Feminicide & Machine Learning: Detecting Gender-based Violence to Strengthen Civil Sector Activism

Catherine D'Ignazio
Data + Feminism Lab, MIT
Cambridge, MA, USA
dignazio@mit.edu

Helena Suárez Val
Feminicidio Uruguay
Montevideo, Uruguay
feminicidio.uruguay@gmail.com

Silvana Fumega
Iniciativa Latinoamericana por los Datos Abiertos (ILDA)
Buenos Aires, Argentina
silvana@idatosabiertos.org

Harini Suresh
Data + Feminism Lab & CSAIL, MIT
Cambridge, MA, USA
hsuresh@mit.edu

Isadora Cruxên
Data + Feminism Lab & DUSP, MIT
Cambridge, MA, USA
cruxen@mit.edu

Wonyoung So
Data + Feminism Lab, MIT
Cambridge, MA, USA
wso@mit.edu

Ángeles Martinez
Data + Feminism Lab & DUSP, MIT
Cambridge, MA, USA
angelesm@mit.edu

Mariel García-Montes
Data + Feminism Lab, MIT
Cambridge, MA, USA
marielgm@mit.edu

ABSTRACT

Gender-related violence against women and its lethal outcome, feminicide, are a serious problem in Latin America and the Caribbean (LAC), as they are in the rest of the world. Although governments have passed legislation criminalizing feminicide, these laws have not been accompanied by relevant policy nor by robust data collection that measures the scope and scale of the problem. Drawing from Data Feminism, we situate feminicide data as "missing data" and describe the work of activists and civil society organizations who attempt to fill in the gaps by compiling incidents of feminicide from news reports. Activists doing this work face challenges: lack of time and financial resources, difficulties in accessing official data, and the mental health burden of reading about violent deaths of women. In this article, we describe our work-in-progress on a participatory action research project designed to help sustain activist efforts to collect feminicide data by partially automating detection using machine learning. We created and labeled a data set for identifying feminicide from media reports and trained a model using this data. The accuracy of the model on our test data set was 81.1%, which shows promise for reducing the labor required to identify and log feminicides, among other potential benefits. We outline our ideas for deploying this model as part of an interactive feminicide notification system, drawing from a co-design process with activists. In the discussion, we raise on-going questions and unresolved tensions that we continue to reflect on while undertaking this work.

MD4SG'20, August 17–18, 2020

ACM ISBN 978-1-4503-2138-9.

DOI: 10.1145/1235

INTRODUCTION

Feminicide & Data Collection

Feminicide (or femicide) is the misogynous and gender-related killing of women by men. The term femicide emerges from the feminist work of Radford & Russell [55], where they define femicide as a form of sexual violence that includes a variety of verbal, physical, visual, and sexual abuse. Based on this work, the term feminicidio (feminicide) was later proposed by Lagarde y de Los Ríos [43] to situate killings of women as part of the entirety of violations of women's human rights[1]. Unlike femicide, the concept of feminicide, widely adopted in Latin America, incorporates impunity that refers to the exemption from punishment. Following Lagarde y de Los Ríos [43], we use the term feminicide to frame the role of the state in enabling violence against women through either omission, negligence or complicity, including neglecting data collection on the issue [43, 61].

Feminicide is a serious problem in the Americas, as it is in the rest of the world. In the United States, around three women

[1]This term was re-introduced as feminicide, from Spanish-speaking Latin American scholarship into English-speaking academia via the edited collection, *Terrorizing Women* [19].

are killed every day by current or former partners [46]. In Latin America, every two hours a woman is killed in the region in incidents related to their gender, according to 2018 data initially compiled by the Economic Commission for LAC (ECLAC) [10]. Intense and persistent activism on the issue from feminist and women's movements across the region has been fundamental to raising awareness and promoting policy change. In Argentina, for example, the powerful demonstrations held by the "Ni una menos" ("Not a [woman] less") movement in 2015 brought worldwide attention to the issue of feminicides and contributed to the formulation of a national plan for the eradication of violence against women [6]. Across LAC, all countries have now passed legislation criminalizing feminicide or femicide with the exception of Cuba and Haiti [11, p. 34]. Feminicide is now a more visible problem in most countries in the region–one that challenges societies and governments to take action.

However, this visibility has not found a correlation in the improvement of the official registration systems of these incidents. Data about violence against women are often neglected by public authorities: they have poor quality, they are difficult to obtain, they are underreported due to stigma and victim-blaming, and they are often contested. The monitoring of feminicides is particularly challenging. Existing official records tend to be incomplete or infrequently updated and often do not allow for comprehending the context for each incident, making it difficult to classify murders of women as feminicides [24, 2]. As a result, policymakers and society are unable to make systematic use of resources and strategies to more effectively tackle this on-going problem. As ECLAC and UN Women state, legal and criminal reforms need to be accompanied by better tools for data collection and monitoring of feminicides [51, 50, 33].

When the state fails to take action through its absence or neglect, civil society organizations and individuals step in to fill in the blanks. As Alice Driver notes in the case of Mexico, "the most accurate records of feminicide are still kept by individuals, researchers, and journalists, rather than by the police or a state or federal institution" [15, p. 7]. The reality is much the same in other countries. In Uruguay, for example, Helena Suárez Val has maintained a Google Map of feminicide cases across the country since 2015 [60]. She situates the digital mapping and visualization of feminicides as an "affect amplifier": it mobilizes feminist affects and emotions to effect change.

Our research project sits at the intersection of the challenges of data collection on feminicides and activists' efforts to monitor and document the problem. In this paper, we detail work-in-progress our research team has undertaken to answer the question: *Can partially automating detection of feminicide cases in a particular geographic context aid civil society groups in their monitoring efforts?* The paper stems from a broader collaborative research project designed to explore how technology, and specifically artificial intelligence, can contribute to mitigating some of the hurdles activists face in gathering counterdata about feminicides. Our research team consists of three partners representing different perspectives on and approaches to the issue. The Data + Feminism Lab, based at MIT, uses data and computational methods to work towards gender and racial equity. The Latin America Open Data Initiative (Iniciativa Latinoamericana para Datos Abiertos, ILDA) has developed a regional data standard for collecting feminicides data in Latin America and the Caribbean, and has been working with national governments in order to understand how they may work towards implementing the standard [21, 22, 20, 36, 37]. Finally, Feminicidio Uruguay is the aforementioned project started by activist and researcher Helena Suárez Val to record and geolocate cases of feminicide in the country [1].

The paper proceeds as follows. We begin by grounding our research effort within the framework of Data Feminism [12], which seeks to both make visible and contribute to dismantling power asymmetries undergirding data collection and analysis. We then provide additional context on the problem of feminicide monitoring and explain our methodological approach. Since this work represents a sociotechnical contribution, we have situated technology development within a participatory action research framework (PAR). The remainder of the paper describes preliminary results obtained from creating test and training data sets and using these to train a classifier to predict the likelihood of feminicide from parsing news articles, and outlines human-computer interaction ideas that we are exploring for implementing a notification system to integrate into activists' workflows. Finally, in the Discussion section, we raise open questions and tensions that have arisen in undertaking this work.

Data Feminism

The theoretical grounding for this work draws from Data Feminism, a set of seven principles that describe how to integrate an intersectional feminist lens into data science [12]. The first two principles of Data Feminism are examine power and challenge power, and are particularly relevant for characterizing the current sociotechnical situation of feminicide data. In these principles, the term "power" is drawing from Patricia Hill Collins' "matrix of domination" – a model for how oppression works at multiple scales [7] – and is used to denote "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 disadvantages—because those same systems were not designed by them or with people like them in mind." [12, p. 24].

In describing the first principle of Data Feminism – examine power – the authors make the case that bias is not only found in data sets after collection. Imbalances of power in the collection environment influence what data is collected (or not collected) in the first place. As they explain, feminicide data can be understood as a case of "missing data": data that are neglected to be prioritized, collected, and maintained, despite their relevance to the well-being of large groups of people. As Alice Driver observes, there is a disquieting parallel here with the disappearances that often mark the murder of women: missing bodies are accompanied by unrecorded violence [15]. Indeed, Lagarde y De Los Ríos' precise definition of feminicide [43] points to the neglect and complicity of the state in enabling these murders, which is made manifest through not measuring the problem. Missing data, in fact, disproportionately have to do with minoritized groups: women, people of color, ethnic minorities, and Indigenous populations. Examples of gendered and racialized missing data abound [2]. In the U.S. context, for example, these include citizens killed by police and maternal mortality statistics.

When the state and other institutions fail to collect important information, activists and civil sector organizations step into those data gaps and collect their own data in order to challenge power, the second principle of Data Feminism. This can be understood as counterdata collection and may be undertaken by activists, journalists, non-profit organizations, citizens, and other groups [27, 9]. Feminicide data collection efforts by activists and civil society groups represent precisely such counterdata collection.

Feminicide Monitoring as Counterdata Collection

As a way to counter the impunity surrounding feminicide, feminist and women activists in Latin America and the United States have taken upon themselves to do the work that states have neglected, collecting counterdata about cases of feminicide from news reports and other independent sources [35, 60]. These mapping and monitoring efforts both highlight and attempt to overcome the inadequacy of official statistics. We have existing connections with women-led organizations and allies which are leading work in this field, and we are in the process of undertaking a qualitative research study to understand why and how these groups monitor feminicides. These organizations provide a crucial denunciation, accountability and transparency function, keeping the issue in the public eye, providing statistics for media and civil society, and pressing governments for structural change. One of the partners on our project, ILDA, has developed a regional data standard for improving official data about feminicide. The standard, geared towards guiding official government data collection, proposes 67 fields that institutions should be collecting about each case [36]. Together, the work of these organizations demonstrates the need to improve both official and activist data practices, while making feminicide a more visible problem.

However, the work of activists and civil society organizations is beset by lack of time and resources, difficulties in accessing official data, and the challenges of interpreting cases from media reports that are often sensationalized, inaccurate or incomplete [15, 17, 44, 2]. The vast majority of the work of monitoring feminicide is unremunerated volunteer work, which includes technical as well as emotional labor. Moreover, due to the various nuances– in legal frameworks but also in feminist activist and academic definitions– of the concept of feminicide adopted in these works and the differences in data collection and visualisation methodologies, it is difficult to compare the data collected through various activist initiatives across different countries [63, 56, 21, 15]. Likewise, these activities are not always well-connected to each other and to regional, national and international civil society organizations who are also trying to work on the issue. Beyond these challenges, manually logging femicides information takes a toll on activists' mental health and has led to physical security risks for some of them.

Our research project aims to help address some of these challenges by collaboratively developing a machine learning model to parse news articles and classify them as feminicides, as well as design an interactive application to notify activists about probable feminicides in the regions they monitor. One of the core objectives of our project is to support and sustain the existing monitoring work of activists and civil society organizations. As geographer Sarah Elwood describes in relation to digital payments systems used by street newspaper vendors, "these turns to digital payment and locative interfaces are aimed at *sustaining* street papers themselves, *not replacing* them, at base a rejection of the disruption logics of technocapitalism." (italics from us) [16]. Thus, it should be clear that the goal is not to replace human labor in order to create an automated system for a faraway central authority to gaze on different countries. Rather, the goal is to make counterdata collection more accessible for newcomers and sustainable for the people who are already doing the work. Crucially, as the Data Feminism approach emphasizes, this requires illuminating and engaging meaningfully with the fact that before or behind the data, there are people counting, documenting, analyzing, and presenting it. Our methods flow from this understanding, for this means we need to be in relationship with the people doing the work to comprehend their experiences, their practices, and their own objectives. This is consistent with both a participatory action research (PAR) strategy–which we discuss below–and the Data Feminism principle "embrace pluralism," which builds on feminist theories of knowledge production to advance the idea that feminist objectivity consists of bringing together multiple, partial, situated perspectives [30].

METHODS

Our project builds on a mixed-methods research design. The technical development of the machine learning model and interactive application for feminicide monitoring are anchored on an iterative, qualitative research process organized around two pillars. The first consists of a co-design process involving the MIT research team and researchers and activists from our partner organizations. The co-design strategy aims to engage our partners in different steps of the research process and tool design—from question definition to database building and tool development—so that the final product reflects the collective knowledge, experience, and needs of those directly engaged in the work of monitoring feminicides. This strategy is grounded within a participatory action research (PAR) framework and research ethic that is committed to joint-learning, collaboration, and to the production of knowledge that can support action and social change [28, 5, 25]. Epistemologically and methodologically, PAR aligns with Data Feminism's call for disrupting conventional unilateral and extractive models of

[2]We use the term "minoritized" to describe groups of people who are positioned in opposition to a more powerful social group. While "minority" describes a social group that is comprised of fewer people, minoritized indicates that a social group is actively devalued and oppressed by a dominant group, one that holds more economic, social, and political power. With respect to gender, for example, men constitute the dominant group, while all other genders constitute minoritized groups [29].

scientific research and challenging expert power hierarchies in knowledge production.

The second pillar extends from the first and consists of qualitative interviews with individual activists and organizations—primarily from the United States and Latin America—dedicated to monitoring feminicides. The interviews aim to understand the workflow, data collection process, and conceptual categories through which actors identify and document feminicide. The interviews will help to provide insight into the practical experiences and challenges of doing this kind of work across different contexts and will also contribute to the development of our models and interactive tool. We have developed an interview guide based on exploratory, informal conversations with different organizations and have begun contacting relevant activists and organizations for potential interviews. Our process for selecting interviewees has been a combination of purposive and snowball sampling—based on recommendations from our partners and other connections—with the aim of reaching a broad range of actors involved in feminicide monitoring. We aim to interview up to 20 people within the next two months via Zoom. The interviews were designed to last around 120 minutes and cover personal / organizational background, categorization and data collection process, and reflections on lessons learned and challenges associated with the monitoring work. Additionally, we aim to show the interviewees proposed mockups and early prototypes for the interactive tool to hear their feedback and suggestions for further development. With the permission of the interviewees, the interviews will be recorded and then transcribed and analyzed using grounded theory and emergent coding, looking for themes and patterns in activists' experiences.

In the next section, we introduce our machine learning model and discuss its development to date. Before moving forward, however, we want to situate our approach to the use of automation to support data collection on feminicides. As D'Ignazio and Klein note in Data Feminism, "the process of converting life experience into data always necessarily entails a reduction of that experience" [12, p. 10]. In the context of feminicides, this process of simplification risks turning the murder of women into mere data points and rendering further invisible the life stories and the concrete loss and pain—individual and collective—associated with violence against women and feminicides. In undertaking this project, we are wary of this risk and we ask ourselves—in an ongoing process of collective dialogue and reflection—how we can honor the people and lives behind the data points. Here, we are inspired by McKittrick's provocation, in the context of racial violence, to confront the "uncomfortable mathematics" of Black deaths and engage with numbers in ways that do not repeat or re-inscribe the violence the numbers represent [47]. Perhaps one tentative approach is to see the integration of automation into feminicide data collection and visualization as a form of counter cultural production that seeks to support the memorialization [15] and affective mobilization practices [60] of activists. As Driver observes, "works of cultural production employ various tactics to recover the full narratives of these women and to rescue their bodies from a discourse that reduces women to the sum of their body parts" [15, p. 21].

USING MACHINE LEARNING TO DETECT FEMINICIDE EVENTS FROM NEWS MEDIA

In this section, we describe the dataset creation and model development of a machine learning system to predict the probability of feminicide from a news article, thus allowing activists to prioritize their time and effort on articles that are much more likely to be relevant. While our research spans several countries, here we report our initial results for a model trained in the English language. We are currently working on developing a model in Spanish based on news articles from Uruguay. During our initial conversations with activists about their current processes, we found that the process of reading through articles to find those mentioning a feminicide is both time consuming and mentally taxing. Our initial efforts are aimed at reducing the labor and mental health burden for this particular step in their workflow. Secondarily, partially automating detection of feminicides may support efforts to systematize data collection on feminicides across different contexts and help to inform policy advocacy, though standardizing definitions and taxonomies around feminicide itself brings challenges which we elaborate in the final section of this paper. Finally, partial automation may aid activists and organizations to more easily initiate and sustain a feminicide monitoring effort in a new geographic context where civil society has not previously been monitoring.

Creating and labeling a data set

Since a labeled dataset of news articles indicating feminicide does not exist, our first step was to collect one. Our overall methodology was 1) collecting a set of news articles, 2) labeling news articles with multiple human annotators, and 3) resolving discrepancies with expert consultation. These steps are described in more detail below.

We collected the initial pool of articles using the open source media analysis platform Media Cloud (mediacloud.org), beginning with English-language articles from sources based in the US. To ensure that the type and distribution of articles was close to what would be seen in activists' normal workflow, we utilized a search query currently used by Feminicidio Uruguay, translated into English:

> *(murder OR homicide OR femicide OR feminicide OR murdered OR dead OR death OR killed OR murdered OR shot OR stabbed OR struck OR strangled OR "lifeless") AND (woman OR girl OR "a young woman" OR "a teenage girl" OR "a girl" OR "body of a woman" OR prostitute OR "sex worker")*

This initial set was filtered to narrow the results of stories to those related to murders in the year 2019, resulting in 338,387 articles. These were then sampled extracting around 40-50 articles per month, and filtered to remove results that were audio/TV transcripts, international news, sources of disinformation or unavailable (404 error), resulting in 393 articles.

The articles were labeled by three people independently using ILDA's definition of feminicide [35]. ILDA's guide to standardising the collection of feminicide data follows the UN's Latin American Model Protocol for the Investigation of gender-related killings of women (femicide/feminicide)

	Category	*Definition*
1	Feminicide	Feminicide refers to the violent death of one or more women based on gendered motivation. Following definition from ILDA [35].
2	Feminicide: stigmatized occupation	A feminicide where the victim is a woman in a stigmatized occupation (e.g., sex worker).
3	Feminicide: transfeminicide	A feminicide where the victim is transgender or gender non-conforming.
4	Feminicide attempt	An attempt at feminicide (i.e., extreme violence) where the victim survived.
5	Linked feminicide	A murder enacted to hurt a woman, where the person who is killed someone else; e.g., a child or new partner. This may be predetermined or happen in the course of an attack on the woman.
6	Not enough information to determine feminicide	A homicide where there is not enough information to determine whether it is or is not a feminicide.
7	Not feminicide	Any occurrence that does not fall into the previous categories.

Table 1. Definitions of categories for labeling feminicides.

Category	*Count*
Feminicide	147
Feminicide: stigmatized occupation	1
Feminicide: transfeminicide	26
Feminicide attempt	29
Linked feminicide	4
Not enough information to determine feminicide	33
Not feminicide	161
TOTAL	399

Table 2. Breakdown of labels present in dataset.

(2014). According to this definition, there are four indicators that outline contextual circumntances of feminicide: (i) a previous relationship between the victim and the perpetrator, (ii) sexual violence, (iii) violence aggravation and (iv) existence of prior complaints. Based on an initial screening of the articles, seven categories that emerge from the main definition were determined; these are detailed in Table 1. Labeling data with more granular categories (e.g., transfeminicide or linked feminicide) allows us to both keep track of how many of these instances are contained in the dataset and to monitor the performance of future models on these less-represented cases.

23.2% of cases contained a discrepancy between the three labelers. The research team convened and reviewed the cases together, as well as double-checked hard-to-label cases against the definition of feminicide from ILDA. A second iteration through these cases to resolve discrepancies left 6.2% still containing disagreement. The majority of these cases were those where one or two labelers indicated that there was not enough information to determine feminicide. This set is representative of a gray area where, even given the ILDA guidelines defining feminicide, it can be difficult to align different interpretations. For these cases, we obtained an additional expert label and included those where three or all four out of four labelers agreed in the final dataset.

Because our dataset did not include any cases of transfeminicide, and the language used to describe these cases can differ from that used in other feminicide cases, we manually collected and added a set of 26 more transfeminicide cases because we wanted to ensure that our classifier could recognize transfeminicide as feminicide. This is both consistent with intersectional feminist thought – the political category of "women" includes cis and trans women – and with Feminicidio Uruguay's data collection practices [59, 62]. The makeup of the final dataset is described in Table 2. It is important to note that we also have very few cases of feminicide of women from stigmatized occupations and linked feminicide, so the resulting classifier would not be appropriate to use for a project monitoring feminicides of sex workers, for example.

Training a Feminicide Detection Model

The full text of each article was cleaned to remove symbols and numbers, under the assumption that those are not yielding useful information about the topic. Each article was vectorized using TF-IDF [39]. We excluded default English stop words as well as date-related words (e.g., month and weekday names) that are prevalent in news articles. The most and least frequent 5% of words were also excluded [57]. This resulted in a 736 word vocabulary.

For model training, categories 2 through 5 in Table 1 were labeled as 'Feminicide' – we do not currently have enough data for these more granular categories to train a multiclass model, and initial conversation with activists suggested that in practice they would want these cases to be flagged as well. How to treat articles without enough information to determine feminicide (category 6) is an open question. In the current iteration, they are excluded during training, but analyzed during testing to explore the performance of the model on these cases.

The data were split into a training and testing set with an 80:20 ratio, resulting in 294 training examples and 74 test examples. The overall percentage of positive (feminicide) articles was 56.8%. We trained a multinomial naive Bayes model using scikit-learn [54] to predict the probability of feminicide from an article.

Model Results

Model performance in terms of various metrics are in Table 3; errors are further detailed in Table 4. Most errors are false positives, as is reflected in the high recall score. In this particular application, high recall is most important: it is better to flag a few extra cases than to miss some.

Metric	Score
Accuracy	0.81
Precision	0.78
Recall	0.93
F1	0.85
AUC	0.88

Table 3. Breakdown of model performance in terms of various metrics.

The breakdown of classification accuracy per category is in Table 5. This more granular evaluation is currently limited by the size of our test set and under-representation of certain categories, but we plan to continue monitoring it as we add in more data. The lowest performance is in the "Not Feminicide" category because most errors are false positives; as noted previously, though, false positives are in general preferable to false negatives.

	Predicted: Feminicide	**Predicted:** Not Feminicide
Actual: Feminicide	39	3
Actual: Not Feminicide	11	21

Table 4. Confusion matrix for model predictions.

When computing accuracy, we derived binary predictions from predicted probabilities using a threshold of 0.5. However, in the interactive application, activists will likely see a list of articles ordered by probability. Therefore, it's useful to look at how valuable such a sorting mechanism would be. We sorted the 74 articles from our test set (42 feminicide and 32 not feminicide), along with the 33 articles with not enough information, by the predicted probability of feminicide output by the model. Then, we looked at the breakdown of articles that would be returned as the number of total articles increased. As seen in Figure 1, actual feminicide articles make up almost all of the first 20 articles and the majority of the first 40. As more articles with lower predicted probabilities are introduced, articles with not enough information increase first, followed finally by the non-feminicide articles.

Category	Num test examples	Accuracy
Feminicide	29	100%
Feminicide: stigmatized occupation	0	n/a
Feminicide: transfeminicide	6	83.3%
Feminicide Attempt	6	83.3%
Linked Feminicide	1	0%
Not Feminicide	32	65.6%

Table 5. Breakdown of model classification accuracy on specific categories.

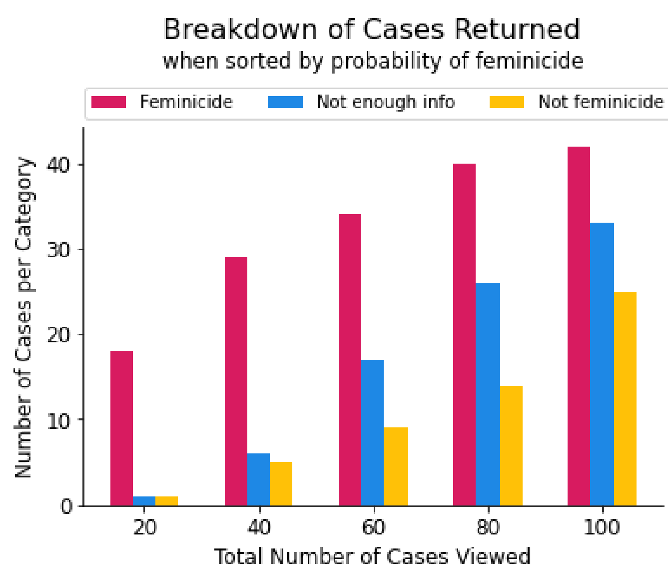

Figure 1. The number of articles returned per category (feminicide, not enough information, not feminicide) as the total number of articles viewed increases. Articles are first sorted by the predicted probability of feminicide output by the model. The first 40 articles are almost all true feminicides, and as the number of articles increases, first articles without enough information and then non-feminicide articles appear.

We emphasize that the goal of any resulting tool is not to replace the process of recording feminicides, but to help make the process less emotionally-taxing and time-consuming for activist groups who are already doing this work. Currently, activists view sets of articles returned by search queries, but these are not sorted in any way. One activist we spoke to stated that for every article actually describing a feminicide, she might look through 20 non-feminicide articles. Simply sorting articles by their predicted probability of feminicide can drastically improve this ratio and serves as a proof-of-concept that such a system could reduce the burden of labor for activists in this space.

System Architecture

Beyond an effective model, it is crucial that the overarching system and interface is well-integrated into activists' current workflow of finding, managing, and extracting information from news articles. Our initial goal is to enhance activists' current workflow by providing a list of articles sorted by likelihood of feminicide.

Figure 2 shows the system architecture. The system uses Media Cloud, an open-source platform for media analysis. When an activist creates their account with their preferred language, a pre-trained, language-specific model is allocated to the user. As a configuration setting, they can define a specific query to match their needs. Along with the model, the configuration is sent to the Media Cloud server, which pulls the articles according to the query defined in the configuration. The machine learning model, subsequently, predicts the probability of feminicide for each article, and the results are returned to the Feminicides server to be displayed to the user. Each user can define their own search query and is assigned their own model, allowing the overarching system to support activists across different contexts. It also provides the capability for real-time updating of model parameters based on user feedback.

Extensions

The ability to prioritize articles by likelihood of feminicide is an important way to make this work less intensive for activists. However, based on conversations with activist groups, there are several other extensions to the system that could further aid in this work:

Returning important entities

After deciding that an article contains a feminicide, activists must extract relevant entities (e.g., name of victim, date of incident, etc) to record in their database. An extension to our system would be to automatically suggest these entities. Media Cloud extracts person and place names by default, and other open source entity extraction tools [18] are able to extract other important terms like dates or ages. These entities could then be fetched and displayed along with the probability of feminicide.

Grouping articles describing the same case together

When looking at articles, activists often do further searches for more articles describing the same case to find additional information. To this end, our system could be extended to return groups of related articles instead of individual articles. In practice, this type of grouping can be implemented using a community detection algorithm, where shared entities are used as edges between articles [48].

Real-time updating of the model with user feedback

The model is trained on a limited set of articles, and as more are seen, performance could be improved by continuing to update the model with user feedback. Moreover, users in different contexts might have different goals for their data collection or might wish to treat "gray area" cases differently. For example, some advocacy groups may focus specifically on monitoring feminicide for trans populations, sex workers or women killed in encounters with police. Rather than use a one-size-fits-all model, it might be worthwhile to allow users to further fine-tune models to their needs.

DISCUSSION

In this section, we raise open questions that have arisen in undertaking this work. Because this is work in progress, we do not yet offer answers but rather ways that we are navigating these questions while engaging in a design process.

The limitations of monitoring feminicide from media reports

Monitoring feminicide through media reports represents a potentially scalable way to track the problem, and governments themselves will sometimes make use of media reports (or even activist records based on media monitoring) as a way of complementing and validating their official data [23]. Yet, scholars and activists acknowledge the limitations of deriving information from the media due to journalistic biases [17, 64]. For example, it has been demonstrated that the deaths of white, middle- to upper-class women are more widely reported than deaths of other women. These media stereotypes render certain women "as valuable 'front-page victims', while dismissing others as disposable." [58]. Monitoring from media reports, therefore, has the potential of amplifying existing race, class and gender biases rather than fully resolving the problem of missing data. Indeed, the organization Sovereign Bodies focuses on monitoring missing and murdered Indigenous women but does not not rely on media reports because of consistent lack of reporting on Indigenous victims and the organization's "families first" approach to the issue. They discover feminicides through Indigenous networks and often from families directly [38]. At the moment, we are navigating these questions by situating civil society monitoring efforts as a partial solution for holding institutions accountable amidst deeply asymmetric power relations. We are additionally mitigating media bias by augmenting our data set in key categories such as transfeminicides and feminicides of women in stigmatized occupations. These cases are less likely to surface in media reports, but should not be overlooked as feminicide. As researchers and designers, we ask – how can we remember and ethically disclose that feminicide data from media reports are always partial, incomplete, and still themselves deeply influenced by inequality?

Differences across language and geography

Because we hope to aid activist organizations across global contexts, as well as potentially aid governmental organizations in the future, we are constructing training and testing datasets spanning different languages. How widely a particular pre-trained, language-specific model can be used across different

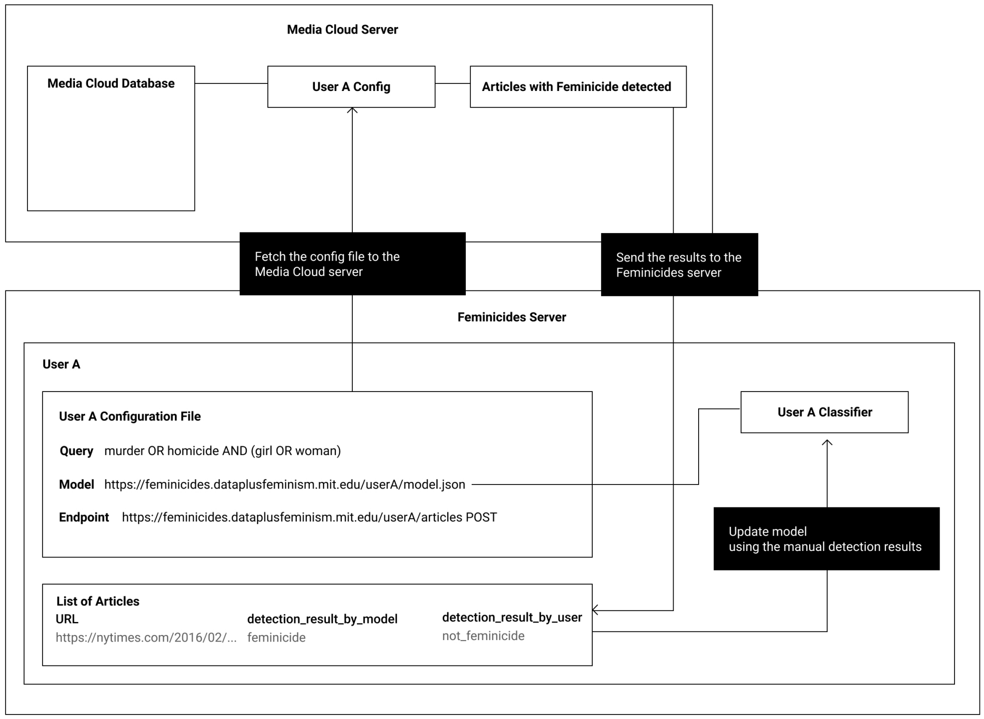

Figure 2. Diagram of system architecture, integrating a local server with Media Cloud.

countries or contexts, however, is an open question. For example, is the way that Peruvian media discusses feminicide fundamentally different from Uruguay? We can begin to investigate this question both qualitatively, by analyzing how media discussion of feminicide differs across countries, and quantitatively, using methods to detect distributional shifts in datastreams [14]. If it appears that context-specific models are better suited to activist goals than a single language-specific model, our proposed system architecture enables fine-tuning a shared, pre-trained model for different users. Such methods draw from research in fine-tuning language classification models to different text corpora [34], model personalization [41], and online learning [13, 26]. They are also aligned with the Data Feminism principle "consider context" which honors and values local context, experiential knowledge and situated knowledge [30, 12].

Unsettling standardization and scale

Automated technology such as that which we are developing would appear to promise standardized and scalable data, but this begs the question, "Why should feminicide data be gathered in a standardized way across geographies, using a standardized definition, to enable regional and global comparisons?" While standardization and scale are often framed as unquestioned normative "good things," an intersectional feminist approach to data involves troubling these assumptions. As geographer Sarah Elwood writes, "Digital principles of standardization, scope/scaling and speed/volume mirror the foundational logics of capitalism to position digital praxes as always already an inherent good within the economism that defines contemporary North American social settlements." [16] Standardization and scale typically translate to greater efficiencies and greater profits – but for whom? Suárez Val [61] has written about how the diverse taxonomies of feminicide created by activists represent not only different definitions of the phenomenon but also different ontological politics - the data categories themselves bring forth (or suppress) different visions of feminicide, different perspectives on who is included, and different perspectives on who is accountable. This points to the importance of coalition building and inclusion for the development of any data standard, which ILDA has written about [21]. Effecting change in the structural or disciplinary domain of the matrix of domination [7] involves shifting law, policy and its implementation, which require comparable data and standardization. For example, the UN has called for establishing a feminicide watch in every country [32]. If we aim to develop high-quality regionally and globally comparable feminicides data, for whom is that data and what will those people do (or continue to not do, since data don't necessitate action on the part of people in power) with them? What local context is lost in the process of standardizing? What is gained by broadening the scale?

Participatory methods for machine learning – with whom?

Machine learning systems have been critiqued for being black boxes [53], meaning they are designed by experts and it is hard to interrogate how and why automated decisions are made, as well as contest those decisions [42]. When ML systems are deployed in the social and political realm, in which historical and present inequalities influence data collection and processing, this opacity combined with flawed data have the potential for producing a "weapon of math destruction," as Cathy O'Neil has famously termed it: a pernicious feedback loop in which existing inequalities are amplified and reinforced [52]. While a number of solutions have been posed for addressing these drawbacks, our work focuses on the value of participatory methods during all phases of the project, from conception to implementation. Drawing from the field of Human-Computer Interaction, there is increasing attention on how technical systems may integrate participatory methods [65]. For example, in Katell & Young et al [40], the team drew from PAR and participatory design methods to work with community organizers and create an Algorithmic Equity Toolkit to help policymakers interrogate automated systems. Similar to our above questions about standardization and scale, the authors find that their co-design partners valued localized resources rather than a toolkit that could be used "anywhere", and that partners found great value in non-technical measures for addressing algorithmic fairness. This is aligned with Kulynych et al's assertion that a focus on algorithms is insufficient for addressing inequitable outcomes of systems [42].

As Katell & Young et al [40] note, working in a model of participatory action research is time- and relationship-intensive, so everybody that may possibly use a system cannot be included. Here we pause to ask a feminist "who question" [12, 49] – who is the community and whose perspectives get priority in a participatory process? For our project on feminicide data, we could have sought out international NGOs and global development organizations as our partners. Drawing from Data Feminism, which requires a power analysis of the data collection environment and places value on embodied, situated knowledge, we chose instead to focus on working with activists working explicitly in a counterdata collection model. This is consistent with feminist HCI's commitment to the "marginalized user" [3, 4] and to the principle advocated by a variety of design methodologies to center the perspectives of the people most impacted by the issue [8, 45, 31].

CONCLUSION

In this paper we have outlined our team's work to date on a system that partially automates the detection of feminicides for civil society organizations and activists who are monitoring the phenomenon. Feminicide is the lethal outcome of gender-related violence against women. While countries across the Americas have made important steps towards passing legislation to criminalize feminicide, new laws have, in most cases, not been accompanied by sufficient policy nor data collection that adequately measures the scope of the problem. Drawing from Data Feminism, we characterize this as a case of "missing data." In the face of insufficient institutional action, lack of or insufficient action, activists and civil society groups are working to collect and monitor cases of feminicide from media reports. Our work is conducted with feminist participatory action research methods in partnership with these groups. We describe a machine learning classifier and interactive application that may help activists more quickly identify cases of feminicide, reduce their labor, and possibly work towards standardizing definitions and taxonomies in the future. Much work remains to be done to develop the system and verify that it contributes to sustaining the existing monitoring work in various locales. In our discussion, we outline four areas of active reflection and engagement for our team: the limitations of detecting feminicides from media reports, differences across language and geography, unsettling standardization and scale, and participatory methods for machine learning.

REFERENCES

[1] 2020. Feminicidio Uruguay. (30 6 2020). https://sites.google.com/view/feminicidiouruguay/inicio

[2] Artigo 19. 2018. Dados Sobre Feminicídio no Brasil. (2018).

[3] Shaowen Bardzell. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2010) *(CHI '10)*. ACM, 1301–1310. DOI: http://dx.doi.org/10.1145/1753326.1753521

[4] Shaowen Bardzell and Jeffrey Bardzell. Towards a Feminist HCI Methodology: Social Science, Feminism, and HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011) *(CHI '11)*. ACM, 675–684. DOI: http://dx.doi.org/10.1145/1978942.1979041

[5] Caitlin Cahill, Farhana Sultana, and Rachel Pain. 2007. Participatory ethics: politics, practices, institutions. *ACME: An International Journal for Critical Geographies* 6, 3 (2007), 304–318.

[6] Laura Rodriguez Claros. 2015. Femicide in Argentina. *Women Across Frontiers Magazine* 2 (2015). https://wafmag.org/2015/11/femicide-in-argentina/

[7] Patricia Hill Collins. 2002. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Routledge.

[8] Sasha Costanza-Chock. *Design justice: Community-led practices to build the worlds we need*. MIT Press.

[9] Morgan Currie, Britt S Paris, Irene Pasquetto, and Jennifer Pierre. 2016. The conundrum of police officer-involved homicides: Counter-data in Los Angeles County. 3, 2 (2016), 2053951716663566. DOI: http://dx.doi.org/10.1177/2053951716663566 Publisher: SAGE Publications Ltd.

[10] Observatorio de Igualdad de Género de América Latina y el Caribe. 2018. Feminicidio. (2018). https://oig.cepal.org/es/indicadores/feminicidio Library Catalog: oig.cepal.org.

[11] Alicia Deus and Diana González. 2018. *Análisis de Legislación sobre Femicidio/Feminicidio en América Latina y el Caribe e Insumos para una Ley Modelo*. Technical Report. MESECVI; ONU Mujeres.

[12] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data Feminism*. MIT Press.

[13] Priyanka B. Dongre and Latesh G. Malik. 2014. A review on real time data stream classification and adapting to various concept drift scenarios. In *2014 IEEE International Advance Computing Conference (IACC)*. 533–537. DOI: http://dx.doi.org/10.1109/IAdCC.2014.6779381

[14] Mark Dredze, Tim Oates, and Christine Piatko. 2010. We're Not in Kansas Anymore: Detecting Domain Changes in Streams. (2010), 11.

[15] Alice Driver. *More or less dead : feminicide, haunting, and the ethics of representation in Mexico*. The University of Arizona Press.

[16] Sarah Elwood. 2020. Digital geographies, feminist relationality, Black and queer code studies: Thriving otherwise. *Progress in Human Geography* (2020), 030913251989973. DOI: http://dx.doi.org/10.1177/0309132519899733

[17] Sarah England. 2018. *Writing terror on the bodies of women: media coverage of violence against women in Guatemala*. Lexington Books. OCLC: on1040183337.

[18] Jenny Rose Finkel, Trond Grenager, and Christopher D Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*. 363–370.

[19] Rosa-Linda Fregoso and Cynthia Bejarano. 2010. *Terrorizing Women: Feminicide in the Americas*. Duke University Press, Durham; London. Google-Books-ID: XFZc9upBBk8C.

[20] Silvana Fumega. 2019a. Open data, gender and violence in Latin America. *European Public Mosaic* 7 (Feb. 2019), 36–47. http://www.gencat.cat/eapc/epum/N7/pdf/EPuM7Fumega.pdf

[21] Silvana Fumega. 2019b. Standardisation of femicide data requires a complex, participatory process, but important lessons are already emerging | LSE Latin America and Caribbean. (Aug. 2019). https://blogs.lse.ac.uk/latamcaribbean/2019/08/08/standardisation-of-femicide-data-requires-a-complex-participatory-process-but-important-lessons-are-already-emerging/ Library Catalog: blogs.lse.ac.uk.

[22] Silvana Fumega. 2020. Tracking Latin America's Other Pandemic: Violence Against Women. (April 2020). https://live-aqimportwithimages.pantheonsite.io/article/tracking-latin-americas-other-pandemic-violence-against-women/ Library Catalog: www.americasquarterly.org.

[23] Victoria Gambetta Sacías and Paula Coraza Ferrari. 2017. Femicidios íntimos en Uruguay. Homicidios a mujeres a manos de (ex) parejas. (2017). https://www.minterior.gub.uy/images/2017/femicidios.pdf

[24] Claudia Garcia-Moreno, Alessandra Guedes, Wendy Knerr, R Jewkes, S Bott, and S Ramsay. 2012. Understanding and addressing violence against women. *World Health Organization, Issue brief No. WHO/RHR/12.37)(S. Ramsay, Ed.)* (2012).

[25] Bev Gatenby and Maria Humphries. 2000. Feminist participatory action research: Methodological and ethical issues. 23, 1 (01 2000), 89–105. DOI: http://dx.doi.org/10.1016/S0277-5395(99)00095-3

[26] Kevin Gimpel, Dipanjan Das, and Noah A Smith. 2010. Distributed asynchronous online learning for natural language processing. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*. 213–222.

[27] Jonathan Gray, Danny Lämmerhirt, and Liliana Bounegru. 2016. Changing what counts: how can citizen-generated and civil society data be used as an advocacy tool to change official data collection? *Available at SSRN 2742871* (2016).

[28] Davydd J Greenwood and Morten Levin. 2006. *Introduction to action research: Social research for social change*. SAGE publications.

[29] Yasmin Gunaratnam. 2003. *Researching'race'and ethnicity: Methods, knowledge and power*. Sage.

[30] Donna Haraway. 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. 14, 3 (1988), 575–599. DOI: http://dx.doi.org/10.2307/3178066

[31] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing community-based collaborative design: Towards more equitable participatory design engagements. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–25.

[32] H Hemblade, A Filip, A Hunt, M Jasser, F Kainz, and others. 2017. Femicide VII: Establishing a femicide watch in every country. (2017).

[33] Helen Hemblade and Helena Gabriel (Eds.). 2019. *Femicide Volume XII: Living Victims of Femicide*. The Academic Council on the United Nations System, Vienna.

[34] Jeremy Howard and Sebastian Ruder. 2018. Universal language model fine-tuning for text classification. *arXiv preprint arXiv:1801.06146* (2018).

[35] ILDA. 2019. Guía para protocolizar procesos de identificación de femicidios para su posterior registro. (03 2019). DOI: http://dx.doi.org/10.5281/ZENODO.1659629

[36] ILDA. 2020a. Estándar de Datos sobre Femicidios ILDA V2 2019. (2020). https://docs.google.com/spreadsheets/d/17fb3XmLCtv42gJj0lhYW43HYU0aqh_ZA77Vyrn6uTy4/edit?usp=embed_facebook

[37] ILDA. 2020b. Standard of Data on Femicides | ILDA. (2020). https://idatosabiertos.org/en/proyectos/estandardatosfemicidios/ Library Catalog: idatosabiertos.org.

[38] National Inquiry into Missing, Murdered Indigenous Women, Girls (Canada), Marion Buller, Michèle Audette, Brian Eyolfson, and Qajaq Robinson. 2019. Reclaiming power and place: the final report of the National Inquiry into Missing and Murdered Indigenous Women and Girls. (2019). https://www.mmiwg-ffada.ca/final-report/

[39] Daniel Jurafsky and James H Martin. 2008. *Speech and Language Processing: An introduction to speech recognition, computational linguistics and natural language processing*. Prentice Hall, Upper Saddle River, NJ, Chapter 6.

[40] Michael Katell, Meg Young, Dharma Dailey, Bernease Herman, Vivian Guetler, Aaron Tam, Corinne Binz, Daniella Raz, and P. M. Krafft. Toward situated interventions for algorithmic equity: lessons from the field. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (2020-01-27). ACM, 45–55. DOI: http://dx.doi.org/10.1145/3351095.3372874

[41] Todd Kulesza, Margaret Burnett, Weng-Keen Wong, and Simone Stumpf. Principles of Explanatory Debugging to Personalize Interactive Machine Learning. In *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI '15* (2015). ACM Press, 126–137. DOI: http://dx.doi.org/10.1145/2678025.2701399

[42] Bogdan Kulynych, Rebekah Overdorf, Carmela Troncoso, and Seda Gürses. POTs: protective optimization technologies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (2020-01-27). ACM, 177–188. DOI: http://dx.doi.org/10.1145/3351095.3372853

[43] Marcela Lagarde y de los Ríos and Charlie Roberts. 2010. Preface: Feminist Keys for Understanding Feminicide: Theoretical, Political, and Legal Construction. In *Terrorizing Women: Feminicide in the Americas*. Duke University Press, xi–xxvi. http://www.jstor.org/stable/j.ctv11smqfd.3

[44] Sonia Madrigal, María Salguero, and Helena Suárez Val. 2019. Monitoring, Recording, and Mapping Feminicide - Experiences from Mexico and Uruguay. In *Femicide Volume XII: Living Victims of Femicide*, Helen Hemblade and Helena Gabriel (Eds.). The Academic Council on the United Nations System, Vienna, 67–73.

[45] Ceasar McDowell and Melissa Yvonne Chinchilla. 2016. 30 Partnering with Communities and Institutions. (2016), 461. Publisher: MIT Press.

[46] Jess McHugh. 2020. Why aren't women in the U.S. also protesting against femicide? (2020). https://www.washingtonpost.com/opinions/2020/03/10/why-arent-women-us-also-protesting-against-femicide/

[47] Katherine McKittrick. 2014. Mathematics Black Life. *The Black Scholar* 44, 2 (2014), 16–28. https://www.jstor.org/stable/10.5816/blackscholar.44.2.0016 Publisher: Taylor & Francis, Ltd.

[48] Elizaveta K Mikhina and Vsevolod I Trifalenkov. 2018. Text clustering as graph community detection. *Procedia computer science* 123 (2018), 271–277.

[49] Michael Muller. 2011. Feminism asks the "Who" questions in HCI. *Interacting with Computers* 23, 5 (2011), 447–449.

[50] Noticias ONU. 2019. Más de 3500 mujeres fueron asesinadas por razones de género en 25 países de América Latina en 2018. (Nov. 2019). https://news.un.org/es/story/2019/11/1465831 Library Catalog: news.un.org.

[51] Observatorio de Igualdad de Género de América Latina y el Caribe. 2019. *La medición del feminicidio o femicidio: desafíos y ruta de fortalecimiento en América Latina y el Caribe*. Technical Report. CEPAL; UN Women. https://oig.cepal.org/sites/default/files/femicidio_web.pdf

[52] Cathy O'neil. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

[53] Frank Pasquale. *The black box society*. Harvard University Press.

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

[55] Jill Radford and Diana E. H. Russell. 1992. *Femicide: The Politics of Woman Killing*. Twayne. Google-Books-ID: 5DCtnQEACAAJ.

[56] Leonie Rauls and Tamar Ziff. 2018. High Rates of Violence Against Women in Latin America Despite Femicide Legislation: Possible Steps Forward. (15 10 2018). https://www.thedialogue.org/blogs/2018/10/high-rates-of-violence-against-women-in-latin-america-despite-f

[57] Hinrich Schütze, Christopher D Manning, and Prabhakar Raghavan. 2008. *Introduction to information retrieval*. Vol. 39. Cambridge University Press Cambridge, Chapter 2.

[58] Sarah Stillman. 2007. 'The missing white girl syndrome': disappeared women and media activism. 15, 3 (2007), 491–502. DOI: http://dx.doi.org/10.1080/13552070701630665

[59] Susan Stryker and Talia M Bettcher. 2016. IntroductionTrans/Feminisms. *TSQ: Transgender Studies Quarterly* 3, 1-2 (2016), 5–14.

[60] Helena Suáres Val. 2017. Affect amplifiers: feminicide, feminist activists, and the politics of counting and mapping gender-related murders of women. (2017).

[61] Helena Suárez Val. 2019a. Data Frames of Feminicide: An ontological reconstruction and critical analysis of activist and government datasets of gender-related murders of women. (08 2019).

[62] Helena Suárez Val. 2019b. Mapping Feminicide. In *MAPPING (IN)JUSTICE SYMPOSIUM*. Symposium conducted at Fordham University, New York City, NY.

[63] UNODOC. 2019. Global Study on Homicide - Gender-related killing of women and girls. (2019). DOI: http://dx.doi.org/10.1023/B:.0000037731.28786.e3

[64] Dawn Wilcox. 2020. personal communication. (2020).

[65] Angela Zhou, David Madras, Inioluwa Deborah Raji, Bogdan Kulynych, Smitha Milli, and Richard Zemel. 2020 (accessed June 29, 2020). *Participatory Approaches to Machine Learning*. https://participatoryml.github.io/.