A Rubric to Identify Misogynistic and Sexist Texts from Software Developer Communications

Sayma Sultana, Jaydeb Sarker, Amiangshu Bosu
{sayma,jaydeb.sarker,amiangshu.bosu}@wayne.edu
Wayne State University
Detroit, Michigan, USA

ABSTRACT

Background: As contemporary software development organizations are dominated by males, occurrences of misogynistic and sexist remarks are abundant in many communities. Such remarks are barriers to promoting diversity and inclusion in the software engineering (SE) domain.

Aims: This study aims to develop a rubric to identify misogynistic remarks and sexist jokes specifically from software developer communications.

Method: We have followed the systematic literature review protocol to identify 10 primary studies that have characterized misogynistic and sexist texts in various domains.

Results: Based on our syntheses of the primary studies, we have developed a rubric to manually identity various categories of misogynistic or sexist remarks. We have also provided SE domain specific examples of those categories.

Conclusions: Our annotation guideline will pave the path towards building automated misogynistic text classifier for the SE domain.

KEYWORDS

misogynistic text, sexist joke, software developer

ACM Reference Format:
Sayma Sultana, Jaydeb Sarker, Amiangshu Bosu. 2021. A Rubric to Identify Misogynistic and Sexist Texts from Software Developer Communications. In *ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '21), October 11–15, 2021, Bari, Italy*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3475716.3484189

1 INTRODUCTION

As contemporary software development organizations are dominated by males, occurrences of misogynistic and sexist remarks are abundant in many communities [28]. Even if the number of women entering into the software industry is increasing, 56% of them leave their tech jobs mid-career [14]. A survey called "Elephant in the Valley" collected experience of women (~200) who are working in software industry for more than 10 years; 84% of them shared that they had been called "too aggressive" in workplace and 60% were harassed sexually [30]. Often women feel alienated due to toxic workplace culture and they either try to fit in there without protesting or quit their jobs silently. Tales of misogynistic culture among many software development organizations repel women to enter the tech industry. Therefore, misogynistic cultures are major barriers to improve women's representation in tech jobs [14]. Similar situations also persist in online platforms for female software developers. Online harassment of female developers came into light specially in 2014 after the 'Gamergate controversy' [11]. The yearlong organized movement started from threatening women from the video game industry via online medium and then degraded into offline [4]. Female developers encounter gender harassment, sexist jokes and misogynistic comments in other online platforms too while sharing their frustrations or experiences. For example, one joke is found in devRant[1], an online community of software developers for sharing their frustrations, connect with others on tech or humor, shares as: *First computer programmer was a lady, that's why the language of computer is difficult to understand.* Sexist jokes allow misogyny or disrespect to women in *disguised and deniable form* and maintain the established status quo in everyday lives.

Both 'Sexism' and 'Misogyny' depict the idea of suppressing women and their opportunities but they differ in key aspects. According to the Oxford dictionary, *sexism* refers to *prejudice, stereotyping, or discrimination, typically against women, on the basis of sex.* When a man assumes a woman would not be a good CEO (Chief Executive Officer) because men are better as leaders, that conveys a sexism [3]. Misogyny is a hostile version of sexism. It expresses sexist attitudes through violent comments or behaviour. Adam *et* al. defines misogyny as *hatred or disdain of women* and *an ideology that reduces women to objects for women ownership, use or abuse.* For instance, because of landing in a job over a man, if a woman is called as 'a series of gendered vulgarities' [3], that is misogyny. Manifestations of such attitudes establish male dominance, encourage violence against women, and degradation of them. In workplace setting, hostility towards female and lack of interventions negatively affect everyone regardless of gender and cause job dissatisfaction [21].

Large-scale FOSS communities, such as Mozilla, OpenStack, Debian, and GNU manage hundreds of projects and generate large volumes of communications among their contributors. Therefore, it is highly time-consuming and infeasible for the project administrators to manually identify misogynistic texts. An automated classifier can help project communities in that regard, which is the ultimate goal of this research. To build an automated classifier, we need to have a labeled dataset aka oracle. However, to build such an oracle, we first need to understand the characteristics of sexist and misogynistic remarks in the Software Engineering (SE) domain. Researchers conducted numerous studies to detect sexism

ESEM '21, October 11–15, 2021, Bari, Italy

ACM ISBN 978-1-4503-8665-4/21/10…$15.00
https://doi.org/10.1145/3475716.3484189

and misogyny in different domains e.g. online communities like twitter [4, 10], workplace [17], advertisements [15] etc. But identifying sexism and misogyny in the SE domain imposes a different challenge. There are domain specific keywords or ideas which seem neutral or make little sense to the general audience but imply jokes or statements about women's personal characteristics, intelligence or women's place in a private sphere. While prior studies focused on detecting different types of sentiment in open source projects [8, 24], misogynistic comment or speech detection has not been explored yet. Moreover, there is no existing guideline to annotate remarks as misogynistic or neutral specifically for the communities of software engineers where numerous examples of misogynistic remarks exist in online communities and forums. This research aims to fill in this gap. Therefore, the goal of this study is to design rubrics to identify remarks and jokes that express misogyny in the SE domain.

On these goals, we have conducted a Systematic Literature Review (SLR) to identify 10 existing studies that have presented guidelines to identify misogynistic remarks or sexist jokes from various domains. We have synthesized those papers to develop a classification scheme. The primary contributions of this study are:

(1) A rubric to identify misogynistic remarks in the SE context.
(2) A rubric to identify sexist jokes in the SE context.
(3) We have made the rubric publicly available at: https://github.com/WSU-SEAL/misogynistic-rubric We plan to maintain it for the community.

In the remainder of this paper, we discuss the related works, detail our research methods, present the results in the Sections 2, 3, and 4 respectively. Section 5 and Section 6 discusses threats to validity and concludes this paper.

2 BACKGROUND

Though sexism is used to be conceptualized as an expression of hostility towards females, Glick and Fiske argued that being a multidimensional construct, sexism can reflect relatively positive feelings towards women despite holding stereotypical beliefs about them. They coined these two types of sexist attitudes as hostile and benevolent sexism [16]. Hostile sexism can be spotted easily but benevolent form can be subtle. However, manifestations of any kind of sexism express masculine dominance and are harmful for women[16, 19]. Subsequent studies identified sexist remarks in online and offline communities using the idea of ambivalent sexism. Sarah *et* al. conducted the introductory study to detect misogyny in online communities like twitter and curated a list of common words among misogynistic texts. This list of words has been used later in automatic identification of sexist posts in twitter [18]. Akshita *et* al. labeled a dataset consisting of tweets and used machine learning models to automatically classify those into three categories: hostile, benevolent, and neutral posts [19]. Being inspired by their study Dylan *et* al. developed models using BiLSTM and attention mechanisms to detect sexist remarks in workplace settings [17]. Subsequent studies also have been conducted to identify misogynistic remark in French [10], Spanish [13], Indian English, Indian Bangla, Hindi [7], Arabic tweets [22]. Furthermore, several studies have been published in identifying misogyny in *Automatic Misogyny Identification* track of IberEval conference[12, 23]. However, most of the prior studies focus on identifying misogyny in tweets posted by common people. In the SE domain, Squire and Gazda collected IRC chats and emails from FOSS project developers where they reported a lot of profane and insulting texts [28]. Moreover, they found gender biased profane jokes ('mom-joke' is commonly used) in developers' communication. However, no study has yet explored automated identification of misogynistic text in software developer communications.

Table 1: Number of papers in result set after each step of Systematic Literature Review(SLR)

Selection process	# of papers
Search (IEEE + Google Scholar + Scopus)	491
Removing duplicates	390
First iteration: Selecting studies based on the title	200
Second iteration: Selecting studies based on the abstract	44
Third iteration: Selecting studies with guidelines to detect misogynistic text or joke	7
After Snowballing	10

3 RESEARCH METHOD

To identify misogynistic and sexist text from the communication of software engineers, we need to develop a codebook. Prior studies that provide taxonomy and guideline for identifying misogynistic speech or jokes, are based on twitter posts or comments from youtube in different languages e.g. English [4], Hindi [7], Arabic [7, 22] etc. To aggregate the knowledge, we conducted a Systematic Literature Review (SLR) on identifying misogynistic text or sexist humor in different languages and domains. We followed the SLR approach proposed by Kitchenham and Charters [20] in this process and documented all the steps in this section.

3.1 Planning for SLR

We define our research goal, protocol, criteria for including or excluding publications and search string in the planning phase. In this study, we limited our search to misogynistic text and sexist humor. So, we aim *to develop a codebook for the taxonomy of misogynistic remarks and sexist jokes and their definitions so that we can unambiguously annotate texts or humors used in software engineering domain as misogynistic or neutral.* Research protocol involves executing SLR on existing studies that present any guideline to identify misogynistic speech or jokes. For this purpose, we find related literature and then filter the selected studies based on inclusion criteria (IC) and exclusion criteria (EC) given below.

- **IC1:** The publication presents any specific guideline for labeling misogynistic text or sexist jokes in any language.
- **IC2:** The study is written in English.
- **IC3:** The paper is published in peer reviewed conferences, journals or workshops.
- **EC1:** The study is an MS or PhD thesis because that might not be peer reviewed or might be published as a paper.
- **EC2:** Any publication or article that is not peer reviewed.

Table 2: Rubrics for identifying misogynistic remark in software engineering

Sexist remark type	Definition	Primary studies
Discredit	slurring over women with no other larger intention	[4, 12, 22]
Stereotyping	Description of women's physical appeal and/or comparisons to narrow standards	[4, 12, 22]
Sexual harassment	To physically assert power over women	[4, 12, 22, 27, 31]
Threats of violence	Intent to physically assert power over women through threats of violence or to intimidate and silence women through threats	[4, 12, 22, 26]
Dominance	To preserve mail control, protect male interest and to exclude women from conversation	[4, 12, 22, 26, 27]
Derailing	To justify abuse, reject male responsibility, and attempt to disrupt the conversation in order to refocus it	[4, 12, 22]
Victim blaming	Blaming the victims for the problems they are facing	[7]
Mixed bias	Gender bias might be mixed with other kinds of biases (like religious or regional)	[7]
Sexual objectification	Evoke sexual imagery and are used for attacking someone	[4, 7, 12, 22]
Damning	Contains prayers to hurt women	[22, 27]

- **EC3:** We can not access the full work or paper or book chapter.
- **EC4:** Any publication that discusses only machine learning models or any approach to automatically detect misogynistic texts.
- **EC5:** Any paper that discusses only people's attitude towards misogyny /sexism or how sexist language affects people or motivation for using sexist language.
- **EC6:** Any study that analyzes the presence of misogyny in any particular book or movie or published article.

Similar to the previous SLR studies, we have followed a hybrid approach to collect relevant studies. The strategy involves searching in Google scholar, Scopus, IEEE database, and Snowballing. We chose Google Scholar because it provides more results than any other search engine databases. We also looked into Scopus and the IEEE database as previous studies [20] suggested these resources which provided solid academic publications. Our searching was limited to the title of the study. We used the following search string to identify relevant publications:

> *(misogynistic OR sexist) AND (detect OR detection OR classification OR identification OR classifying OR identifying OR text OR post OR remark OR speech OR joke OR humor OR language)*

Though we aim to build a rubric for identifying misogynistic posts primarily, we included 'sexist' keyword in our search string because misogyny is a part of sexism and we did not want to exclude any study that provides clues to spot misogynistic speech. While developing the rubric for remarks, we included all types of misogyny and violent forms of sexism and for the rubric of jokes, we considered any type of sexism: violent or non-violent.

3.2 Executing the SLR

We followed seven steps to conduct the research protocols formulated above. It involves searching for articles in selected databases, removing duplicates, selecting primary studies based on the inclusion criteria(IC) and exclusion criteria(EC) and finally snowballing to find out relevant studies. Table 1 depicts the number of publications in the result set after each step of the SLR process.

- **Step 1: Searching.** We manually searched for publications on Google Scholar, IEEE database, and Scopus. All searches were completed by June 16, 2021. We ended up with a total of 491 papers.
- **Step 2: Removing duplicates.** We created a list of the titles with author names of the publications and manually removed all duplicate papers from the collection and found 390 unique papers.
- **Step 3-5: Selecting primary studies.** During the first iteration, we looked over the title of the study and their publication venues. Based on the exclusion criteria, we removed 190 entries from the list. After that, we have gone through the abstracts to filter out publications that are irrelevant to our goal. We also removed all MS or PhD thesis during this step. In total, we removed 156 studies during the second iteration. Finally, we conducted full-text reading on the result set of papers and examined if the papers align with our inclusion criteria. Conducting the third iteration resulted in a total of 7 papers.
- **Step 6: Backward Snowballing.** We followed backward snowballing to look for additional papers. In this step, we investigated the references of the selected publications and examined if any study fits with our inclusion criteria. We continued this step until no more relevant study was found. We found 3 additional studies during this step.
- **Step 7: Forward Snowballing.** We followed forward snowballing, using Google Scholar citation search to check papers that have cited the identified studies. This step did not find any additional study.

The SLR approach resulted in a set of 10 publications containing annotation guidelines for misogynistic speech or sexist humor. Two of our team members divided the list of paper into halves after removing duplicates and selected primary studies following the inclusion and exclusion criteria. Having any doubt, they marked the paper as undecided and conducted follow up discussion to make a decision about including or excluding the paper. Though the SLR approach involves contacting prolific authors to find out additional studies, as there is no existing study to identify misogynistic remarks in software engineering, we skipped this step. We extracted the characteristics of misogynistic text and enlisted those in a google doc. After that, we aggregated all similar categories

and developed the rubric for identifying misogynistic texts in a fine-grained manner.

4 RESULTS

Systematic Literature Review resulted into 10 primary studies that contains taxonomy of misogynistic or sexist remarks or jokes. We have aggregated all types of misogynistic texts in 2, sexist jokes in 2. Two primary studies also provide guidelines for identifying toxic texts or jokes based on the target class. We have presented those in table 4. We found a few relevant examples from prior study [6, 29] that presents dataset for real world communications among developers and posts from devRant [1]. We came up with the remaining examples by ourselves following the definitions of the remark types. We understand that without contextualization content can be perceived in different way. But right now we do not have enough dataset for each types of remarks. We aim to mine real world communication data among developers and use our developed rubric to identify sexist remarks.

4.1 Classification of misogynistic remarks

Anzovino et al. [4] divide misogynistic texts into seven categories while annotating twitter data. This categorization has been used by subsequent studies to annotate English and Spanish tweets [12] and Arabic tweets [22]. Mulki et al. [22] add another category *Damning* with existing framework while labeling Arabic toxic content. We incorporated two other categories in our rubric as *Victim blaming* and *Mixed bias* following the study of Bhattacharya et al. [7].

- **Discredit.** Discredit refers to random mean speech towards women without any specific intention. It not only expresses a negative attitude towards women but also encourages other males to take part in such toxic behaviour.
  I don't think girls in any support could fix that [29].
- **Stereotyping.** Stereotypical language establishes typical gender roles of the victim or the aggressor [7] and makes women feel subordinated. The following example establishes the fact that women are bad at problem solving.,
  Speaker 1: Hey my girl friend solved the problem.
  Speaker 2: That's cool, ha! Girls are going smart these days. [29]
- **Sexual harassment.** Sexual harassment occurs when anyone refers to one's sexuality or sexual orientation aggressively [7]. Such types of misogyny may occur to show power over women or others. For instance,
  *AHAHAHAHA WHAT THE F***!!?? I'd screwed more girls then that when i was 14 for f***s sake!! [29]*
- **Threats of violence.**Threats of violence can occur against homophobic and transgender too [7]. One post express violence against women as such,
  WordPress deserves to be treated the same way women are treated in game of thrones. [1]
- **Dominance.** Prior study suggests that men dominate women to compensate for their inadequate feeling [25]. Such types of behaviour assists to maintain the status quo and results into gender injustice. The following example supports the fact that men are dominant in the development area.
  Speaker 1: What made you abandon your roots of shopping and join the dark side?
  Speaker 2: My roots of shopping? :-p
  Speaker 1: Girls like to shop - don't worry, I like to shop too. [29]
- **Derailing.** Derailment takes place when anyone is called out for social injustice or bad behaviour and they divert the discussion into another issue [2]. For instance, during Flash conference in 2009, Hoss Gifford presented sexual images of men and women [2]. When women complained about those, they received responses like-
  That's how the Flash community is/ women shouldn't be too sensitive or prudish; it was humor. [2]
- **Victim blaming.** Victim blaming is a common phenomena in patriarchal society when women are blamed for being abused or subjected to male violence. Moreover, such types of language or behaviour encourage women to blame themselves and other women and girls.
  If you were being harassed, why didn't you talk to the HR before?
- **Mixed bias.** Sometimes gender bias gets merged with other forms of biases, for example, biases or toxic behaviour based on race, color, religion etc.
  Was that the only black girl available? She's not even that hot. Sorry to make such a deal out of it but it was a rude awakening after I rebooted [29].
- **Sexual objectification.** Objectification or sexual objectification refers to treating women as objects. It involves describing women's physical appeal or attacking them using their sexual imagery. For instance,
  Actually I hoped to see some sexy girls before the release. [29]
- **Damning.** Damning involves prayers that contain death or illness wishes to curse women while praying to God [22].
  Just putting it out there, you deserve all those deaths you are getting [27].

4.2 Classification of sexist jokes

Bemiller et al. [5] explored contents of sexist jokes collected from the internet and categorized those into five classes. We incorporated their classification for identifying sexist humor. Geek Feminism Wiki also mentions a few types of sexist jokes [2] which align with the classes proposed by Bemiller et al. [5]. However, we did not include a few categories from geek Feminism like unprofessional, presumptive jokes [2] that involve humor and might be funny among close friends but not in the workplace. We assume those types of jokes can be annotated using the following categories.

- **Devaluation of personal characteristics.** Humor involving women's physical attributes or personal characteristics is common in the online or offline world. The following example portray women as complicated and spendthrift,
  Why computers are like women: 1. No one but the Creator understands their internal logic. 2. The native language they use to communicate with other computers is incomprehensible to everyone else. 3. Even your smallest mistakes are stored in long-term memory for later retrieval. 4. As soon as you make a commitment to one, you find yourself spending half your paycheck on accessories for it. [1]

Table 3: Rubrics for identifying misogynistic joke in software engineering

Joke type	Definition	Primary Studies
Devaluation of personal characteristics	Jokes about personal attributes of women or lack of certain attributes, e.g. jokes about intelligence, breasts, weight, personal characteristics	[5, 9]
Women's place in the private sphere	Humor about cooking, cleaning, and childcare	[5, 9]
Violence against women	Jokes that are violent against women e.g. suggesting to abuse women, killing or raping women	[2, 5, 6]
Feminist backlash	Jokes that portray women's issues as inconsequential and tend to attack someone personally	[5, 6]
Sexual objectification	Humor that objectifies women by constructing them as objects for sexual pleasure of men or by decompartmentalizing women's bodies	[2, 5, 9]
Excluding and/ or objectifying	Humor that deny presence of women and/ or objectifies women	[2, 6]
Transphobic	Jokes that depend on women having body parts typically possesses by men or men wearing women's clothing	[2]
Cruel	Humor by mocking or belittling another person or their distress	[2]

Table 4: Sexual remarks or jokes based on narration

Type	Definition	Primary studies
Direct/ active	Directly addresses to a woman or a group of women	[10, 12]
Passive/ descriptive	Describe a woman or women in general	[10, 12]
Report	Report of a sexism experience	[10]

- **Women's place in the private sphere.** Historically, women are associated with the private sphere like maintaining home, reproduction, and child-rearing while men are associated with paid work and the public sphere [5]. Jokes related to such type of division of labour devalue the contribution of women and reinforce male dominance. For instance,

  Women astronaut: Because sandwiches won't make themselves up here. [1]

  This category also includes jokes expressing men's dissatisfaction in their marriage.
- **Violence against women.** When women do not go along with the stereotypes established by the patriarchal system, humor is used in order to control them. For example,

  If all women secretly want to be raped, you're not a real woman if you don't want to be raped. But since you always get what you really want, if I haven't been raped, maybe I secretly don't want to be a woman [6].
- **Feminist backlash.** Jokes towards women who break the status quo or foster feminist attitudes fall under this category. For instance,

  Why feminists hates linux ?! Because they can't "women" commands [1].
- **Sexual objectification.** Jokes involving body parts of women or considering women as objects are included in this class. Such types of humor can occur on a personal or institutional level to establish masculine superiority. For example,

  We should invite more women into our community so we can get laid [2].
- **Excluding and/ or objectifying.** These jokes are made up assuming that women are not included in the community or they are not interested in geeky topics or can be treated as objects [2]. For instance,

  Comp Science teacher: What is an object?
  Kid sitting next to me: Women. [1]
- **Transphobic.** Transphobic jokes or remarks include humor or disgust towards people who do not conform to the gender expectation of the society [2]. For instance,

  Tor sticks out as badly as a transgendered Mongolian in the desert[2].
- **Cruel.** Cruel jokes are made up from the distress or humiliation of one person [2].

  If I wanted to kill myself, I would jump down to her IQ level.[3]

4.3 Classification of sexist remarks or jokes based on target

Based on the target, Chiril et al. [10] classified sexist texts into three categories while annotating french tweets. They mentioned that the reporting category may not be considered as sexist when tweets are moderated. We included this category as such types of report can help to identify sexist and misogynistic speech or past events.

- **Direct/ active.** Direct or active sexist speech usually contains a second person pronoun or verb (e.g. You or You are [10]). These messages are sent to a specific person. For instance,

  You are a woman, you must be liking the documentations stuff.
- **Descriptive/ passive.** This type of remarks contain generalizing terms or named entities [10]. There remain many potential receivers for such remarks [12]. For example,

  JavaScript is the ugly girl freshman year with glasses and acne. [1]

[1] https://www.pinterest.com/pin/410460953515312574/

[2] https://twitter.com/k8em0/status/542893488191586304

[3] https://twitter.com/menshumor/status/472799711515004930?lang=en

- **Report.** Reporting involves description of any sexism experience or informing such behaviour. Such posts may contain reporting verbs or quotation marks [10]. For instance,

  Worst 'advice' from a college recruiter: "O you want to major in computer science? Well our school is fantastic for women in comp sci because WHEN they find it too difficult they can easily transition to graphic design. How do you feel about graphic design?" [1]

5 THREATS TO VALIDITY

The problem of identification and classification of misogyny and sexism in modern communication medium is still its outset period and extend many challenges. Our annotation guideline projects only misogynistic texts and sexist humor. We did not incorporate any dimension of sexism that is benevolent or express stereotyping about women. Benevolent sexist remarks also impact women negatively and establish gender bias[16, 19]. For example, *You code better as a girl*. These types of comments may exist during real world communication and are harmful for women. We aim to explore benevolent sexism in future research.

We crafted the annotation guideline based on previous studies where misogynistic remarks were collected from twitter[4, 22], comments from Youtube[7]. We want to use this rubric to classify dataset collected from communications of software engineers. Unique cases of misogyny may arise which are not covered by the above rules. However, our research is the first step towards building a complete taxonomy for misogyny and sexist humor. We plan to use this rubric to manually label a large scale oracle and build an automated classifier to detect misogynistic remarks. During the process of manual labeling, if certain text does not fit into above mentioned categories identified in this research, we plan to amend this rubric on Github.

6 CONCLUSION

This study offers a theory-driven misogyny and sexism detection approach from software developer communications. We conducted a systematic literature review on prior studies that unveils various dimensions of misogyny and sexist humor and aggregating the existing literature, we developed a codebook to label communication dataset from groups of developers. Our proposed annotation guideline can be used to manually identify misogynistic texts from the SE domain and pave the way to build automated classifiers for this domain.

REFERENCES

[1] [n.d.]. *devRant: A fun community for developers to connect over tech or humor.* https://devrant.com

[2] [n.d.]. *Geek Feminism Wiki.*

[3] Kate Bahn Annie McGrew. September 28, 2017. What is misogyny, and why should you care? https://medium.com/@amprog/what-is-misogyny-and-why-should-you-care-90ef7d4174aa. [Online; accessed July 13, 2021].

[4] Maria Anzovino, Elisabetta Fersini1, and Paolo Rosso. [n.d.]. Automatic Identification and Classification of Misogynistic Language on Twitter.

[5] Michelle Bemiller and Rachel Schneider. 2010. It's not just a joke. *Sociological Spectrum* 30 (07 2010), 459–479. https://doi.org/10.1080/02732171003641040

[6] Memo Bergmann. 1986. How Many Feminists Does It Take To Make A Joke? Sexist Humor and What's Wrong With It. *Hypatia* 1, 1 (1986), 63–82. https://doi.org/10.1111/j.1527-2001.1986.tb00522.x

[7] Shiladitya Bhattacharya, Siddharth Singh, Ritesh Kumar, Akanksha Bansal, Akash Bhagat, Yogesh Dawer, Bornini Lahiri, and Atul Ojha. 2020. Developing a Multilingual Annotated Corpus of Misogyny and Aggression.

[8] Kevin Daniel André Carillo, Josianne Marsan, and Bogdan Negoita. 2016. Towards Developing a Theory of Toxicity in the Context of Free/Open Source Software & Peer Production Communities. *SIGOPEN 2016* (2016).

[9] Anastasia Nelladia Cendra, Teresia Dian Triutami, and Barli Bram. 2019. Gender stereotypes depicted in online sexist jokes. *The European Journal of Humour Research* 7, 2 (Jul. 2019), 44–66. https://doi.org/10.7592/EJHR2019.7.2.cendra

[10] Patricia Chiril, Véronique Moriceau, Farah Benamara, Alda Mari, Gloria Origgi, and Marlène Coulomb-Gully. 2020. An Annotated Corpus for Sexism Detection in French Tweets. In *Proceedings of the 12th Language Resources and Evaluation Conference*. European Language Resources Association, Marseille, France, 1397–1403. https://aclanthology.org/2020.lrec-1.175

[11] Lize De Clercq. February 28, 2016. Why are there so few female game developers? (February 28, 2016). http://www.unite-it.eu/profiles/blogs/why-are-so-few-women-developing-video-games

[12] E. Fersini, P. Rosso, and M. Anzovino. [n.d.]. Overview of the Task on Automatic Misogyny Identification at IberEval 2018.

[13] José Antonio García-Díaz, Mar Cánovas-García, Ricardo Colomo-Palacios, and Rafael Valencia-García. 2021. Detecting misogyny in Spanish tweets. An approach based on linguistics features and word embeddings. *Future Generation Computer Systems* 114 (2021), 506–518. https://doi.org/10.1016/j.future.2020.08.032

[14] Eliza gauhar. May 23, 2020. The Anatomy of Misogyny in STEM. https://medium.com/novidem-magazine/the-anatomy-of-misogyny-in-stem-8e4b189a971a. [Online; accessed July 13, 2021].

[15] Rosalind Gill. 2008. Empowerment/Sexism: Figuring Female Sexual Agency in Contemporary Advertising.

[16] Peter Glick and Susan Fiske. 1996. The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism. *Journal of Personality and Social Psychology* 70 (03 1996), 491–512. https://doi.org/10.1037/0022-3514.70.3.491

[17] Dylan Grosz and Patricia Conde-Céspedes. 2020. Automatic Detection of Sexist Statements Commonly Used at the Workplace.

[18] Sarah Hewitt, T. Tiropanis, and C. Bokhove. 2016. The Problem of Identifying Misogynist Language on Twitter (and Other Online Social Spaces). In *Proceedings of the 8th ACM Conference on Web Science* (Hannover, Germany) *(WebSci '16)*. Association for Computing Machinery, New York, NY, USA, 333–335. https://doi.org/10.1145/2908131.2908183

[19] Akshita Jha and Radhika Mamidi. 2017. When does a compliment become sexist? Analysis and classification of ambivalent sexism using Twitter data. https://doi.org/10.18653/v1/W17-2902

[20] Barbara Ann Kitchenham and Stuart Charters. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering.* Technical Report EBSE 2007-001. Keele University and Durham University Joint Report. https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf

[21] Kathi Miner-Rubino and Lilia M. Cortina. [n.d.]. Beyond targets: Consequences of vicarious exposure to misogyny at work. Journal of Applied Psychology, Vol 92(5), Sep 2007, 1254-1269. https://psycnet.apa.org/buy/2007-12832-006

[22] Hala Mulki and Bilal Ghanem. 2021. Let-Mi: An Arabic Levantine Twitter Dataset for Misogynistic Language.

[23] Endang Wahyu Pamungkas, Alessandra Teresa Cignarella, Valerio Basile, and V. Patti. 2018. 14-ExLab@UniTo for AMI at IberEval2018: Exploiting Lexical Knowledge for Detecting Misogyny in English and Spanish Tweets. In *IberEval@SEPLN*.

[24] Rajshakhar Paul, Amiangshu Bosu, and Kazi Zakia Sultana. 2019. Expressions of sentiments during code reviews: Male vs. female. In *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER)*. IEEE, 26–37.

[25] Will Damon PhD. 2003. Dominance, Sexism, and Inadequacy. *Journal of Psychology & Human Sexuality* 14, 4 (2003), 25–45. https://doi.org/10.1300/J056v14n04_02

[26] Mattia Samory, Indira Sen, Julian Kohne, Fabian Floeck, and Claudia Wagner. 2021. "Call me sexist, but...": Revisiting Sexism Detection Using Psychological Scales and Adversarial Samples. arXiv:2004.12764 [cs.CY]

[27] Sima Sharifirad and Stan Matwin. 2019. When a Tweet is Actually Sexist. A more Comprehensive Classification of Different Online Harassment Categories and The Challenges in NLP. *CoRR* abs/1902.10584 (2019). arXiv:1902.10584 http://arxiv.org/abs/1902.10584

[28] Megan Squire and Rebecca Gazda. 2015. FLOSS as a Source for Profanity and Insults: Collecting the Data. In *2015 48th Hawaii International Conference on System Sciences*. IEEE, 5290–5298.

[29] Megan Squire and Rebecca Gazda. 2015. FLOSS as a Source for Profanity and Insults: Collecting the Data. *Proceedings of the Annual Hawaii International Conference on System Sciences* 2015. https://doi.org/10.1109/HICSS.2015.623

[30] Fullstack Staff. [n.d.]. How 7 Tech Companies Are Fighting Workplace Sexism. ([n. d.]). https://www.fullstackacademy.com/blog/tech-companies-fighting-workplace-sexism

[31] M. Strauss. 1989. *Sexist Speech in the Workplace.* Harvard Civil Rights-Civil Liberties Law Review. https://books.google.com/books?id=2TKTnQEACAAJ