Preventing #MeToo: Artificial Intelligence, the Law, and Prophylactics

James P. de Haan
Paul Hastings, LLP, jamesdehaan@paulhastings.com

Recommended Citation

James P. de Haan, *Preventing #MeToo: Artificial Intelligence, the Law, and Prophylactics*, 38 LAW & INEQ. (2020).
Available at: https://scholarship.law.umn.edu/lawineq/vol38/iss2/4

Preventing #MeToo: Artificial Intelligence, the Law, and Prophylactics

James de Haan†

TABLE OF CONTENTS

### I.	"When Harvey Met Sally": A Simple Introduction to a Complicated Problem

"There is power in unity and there is power in numbers. As long as we keep moving like we are moving, the power structure . . . will have to give in."[1]

Once there was a man named Harvey.[2] Harvey runs a successful public relations firm, and recently hired Sally as an intern. Sally was twenty-one years old, fresh out of college, and dreamt of running her own PR firm. Harvey told her she could shadow him to learn the ropes and generally help around the office. Sally was over-the-moon; she grew to see Harvey as a mentor. She happily accepts his friend request on Facebook. And she has no issue following him back on Instagram. She was not even dismayed when he randomly liked a two-year-old photo of her at the beach. She worked with Harvey every day, and he was always very friendly!

Occasionally, her work kept her in the office later than the rest of the staff. On one of those nights, Harvey asked Sally if she wanted to have a drink in his office. She agreed, and they began to talk—about her hopes, dreams, family. Her love life. A couple of hours, and more than a few drinks later, Sally thanked Harvey and turned to go. But as she was leaving, she felt his hand on her shoulder. Then, he pressed against her. And she froze. Harvey asked her to stay. She squeaked out a no. And she ran.[3]

Sexual harassment manifests as an abuse of power.[4] Harassers are not typically sexual deviants or addicts; they want

---

1. Dr. Martin Luther King, Jr., Keep on Moving, Speech at St. Luke's Baptist Church in Birmingham, Ala. (May, 1963).

2. Names, characters, businesses, places, events, locales, and incidents are purely products of the author's imagination. Any resemblance to actual persons, living or dead, or actual events is entirely coincidental.

3. Of course, sexual harassment is rarely this cut-and-dry. While traditional *quid pro quo* harassment is the most striking—and perhaps the most focused on—form, harassment is like a disease. It festers and grows in aberrant conditions. And, rather than being this explicit, can manifest as a general feeling of helplessness against a tirade of sexual jokes, unwanted hugs, and backhanded praise.

4. One study found that the three "forms" of sexual harassment—*quid pro quo*, hostile environment, and harassment by a third party—are simply manifest from different bases of power. Abuse of individual power leads to *quid pro quo*; organization power to hostile environment; and societal power to third-party harassment. Paula M. Popovich & Michael A. Warren, *The Role of Power in Sexual Harassment as a Counterproductive Behavior in Organizations*, 20 HUM. RESOURCE MGMT. REV. 45, 49 (2010).

control.[5] The #MeToo movement is a natural backlash against that power dynamic. Men and women, many of whom feel isolated and alone, form a community of the harassed. Social media amplifies their voices; "#MeToo" collects them together.[6] Though some were surprised at the sheer volume of stories, [7] the disenfranchised have always found power in community.[8]

---

5. Studies have also shown it is those who are chronically *denied* power that exhibit a stronger desire to feel powerful, and thus are more likely to use sexual aggression towards that end. Melissa J. Williams, Deborah H. Gruenfeld, & Lucia E. Guillory, *Sexual Aggression When Power Is New: Effects Acute High Power on Chronically Low-Power Individuals*, 112 J. PERSONALITY & SOC. PSYCHOL. 201, 204–14 (2017). Some analysts hypothesize that is why men in supervisory roles are more likely to harass their female colleagues, rather than women they directly supervise: "harassment [i]s an 'equalizer' against women supervisors, consistent with research showing that harassment is less about sexual desire than about control and domination." McLaughlin et. al, *Sexual Harassment, Workplace Authority, and the Paradox of Power*, 77 AM. SOC. REV. 625, 641 (2012).

6. Alyssa Milano was credited with creating the hashtag when she wrote "[i]f all the women who have been sexually harassed or assaulted wrote 'Me too' as a status, we might give people a sense of the magnitude of the problem." Nadia Khomami, *#MeToo: How a Hashtag Became a Rallying Cry Against Sexual Harassment*, GUARDIAN (Oct. 20, 2017, 1:13 PM), https://www.theguardian.com/world/2017/oct/20/women-worldwide-use-hashtag-metoo-against-sexual-harassment [https://perma.cc/3KBA-B623]. However, activist Tarana Burke came up with the "Me Too" designator back in 2006, which she used to refer to her wider movement to help survivors of sexual harassment and assault. Abby Ohlheiser, *The Woman Behind 'Me Too' Knew the Power of the Phrase when She Created It – 10 Years Ago*, WASH. POST (Oct. 29, 2017, 7:38 AM), https://www.washingtonpost.com/news/the-intersect/wp/2017/10/19/the-woman-behind-me-too-knew-the-power-of-the-phrase-when-she-created-it-10-years-ago/ [https://perma.cc/7NDT-QYJA].

7. It is normally men, rather than women, who are surprised by these stories. Mario Small, a sociologist that studies support networks, believes this is partly due to the fact that sexual harassment is one of those incidents that is "too painful to risk the possibility that our confidant, the person we are very close to, might say or do the wrong thing . . . . [W]omen have reason to be unsure how a close male might respond. Many of us men have been harassed, sexually assaulted, or stalked. But for almost none of us is the experience as pervasive, persistent or routine as it is for women." Mario Small, *What 'Me Too' Can Teach Men Who Are Willing to Listen*, TIME (Oct. 19, 2017), http://time.com/4988137/me-too-men-listen/ [https://perma.cc/2HRN-5TAH].

8. Critical theorists have long recognized the importance of community. When "[t]hose injured by racism and other forms of oppression discover that they are not alone and moreover are part of a legacy of resistance . . . . [t]hey become empowered participants, hearing their own stories and the stories of others, listening to how the arguments against them are framed and learning to make the arguments to defend themselves." Tara J. Yosso, *Whose Culture Has Capital? A Critical Race Theory Discussion of Community Cultural Wealth*, 8 RACE ETHNICITY & EDUC. 69, 75 (2005).

Sexual harassment is a blight on modern society.[9] And it is far too prevalent.[10] GfK, a market-research group, and the University of California San Diego interviewed two-thousand men and women in 2017, and discovered "81% of women and 43% of men reported experiencing some form of sexual harassment and/or assault in their lifetime."[11] Thirty-eight percent of those women—and thirteen percent of the men—were harassed at work.[12]

Yet sexual harassment remains grossly underreported.[13] The Equal Employment Opportunity Commission (EEOC) believes "*approximately* 70% of individuals who experienced harassment" never file a formal complaint or speak with a supervising authority.[14] Perhaps that is because around 75% of employees who speak out reported some form of retaliation.[15] Regardless, the gap between experience and reports—as well as the high prevalence of

---

9. Economists typically agree that sexual harassment presents a net drain on resources, but it also harms people in ways that are difficult to quantify. Lynn Parramore, *#MeToo: The Economic Cost of Sexual Harassment*, INST. FOR NEW ECON. THINKING (Jan. 2018), https://www.ineteconomics.org/research/research-papers/metoo-the-economic-cost-of-sexual-harassment [https:/perma.cc/9MD9-9A4A].

10. These meetings do already exist—albeit post-complaint—and they are rarely enjoyable for those involved. Having a machine, rather than a person, prompt these meetings could embolden the representative. Rather than explain why a co-worker filed a complaint, she simply has to say the program flagged the issue, and she can then empathize and work through the incident with the would-be harasser. As Scalia once wrote "[t]he chances that frail men and women will stand up to their unpleasant duty are greatly increased if they can stand behind the solid shield of a firm, clear principle . . . ." Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175, 1180 (1989). STOP ST. HARASSMENT, NATIONAL STUDY ON SEXUAL HARASSMENT AND ASSAULT 7 (Jan. 2018), http://www.stopstreetharassment.org/wp-content/uploads/2018/01/Full-Report-2018-National-Study-on-Sexual-Harassment-and-Assault.pdf [https://perma.cc/Q4U9-K2JP].

11. *Id.*

12. *Id.* at 8. A study from 2015 with a much smaller sample size reached similar conclusions. Alanna Vagianos, *1 in 3 Women Has Been Sexually Harassed at Work, According to Survey*, HUFFINGTON POST (Feb. 19, 2015, 4:34 PM), https://www.huffpost.com/entry/1-in-3-women-sexually-harassed-work-cosmopolitan_n_6713814 [https://perma.cc/6N6Y-PJ7Y].

13. CHAI R. FELDBLUM & VICTORIA A. LIPNIC, EQUAL EMP'T OPPORTUNITY COMM'N, SELECT TASKFORCE ON THE STUDY OF HARASSMENT IN THE WORKPLACE 16 (June 2016), https://www.eeoc.gov/eeoc/task_force/harassment/upload/report.pdf [https: perma.cc/V8B3-9E6H].

14. *Id.*

15. Lilia M. Cortina & Vicki J. Magley, *Raising Voice, Risking Retaliation: Events Following Interpersonal Mistreatment in the Workplace*, 8 J. OCCUPATIONAL HEALTH PSYCHOL. 247, 255 (2003). Interestingly, the authors of the study expanded retaliation beyond the well-studied realm of tangible workplace actions to encompass social retaliation, including "harassment, name-calling, ostracism, blame, threats, or the 'silent treatment.'" *Id.* at 248.

retaliation—is evidence of obvious problems in sexual harassment law.

But what could employers do? Around 70% of employers already offer training; 98% of companies have adopted a written sexual harassment policy.[16] Yet harassment continues, even in companies with some of the clearest and most confidential reporting.[17] Often humans are the root of failure—"sexual harassment policies . . . are only as good as the managers who implement them and are responsible for making sure there is broad compliance."[18]

This paper explores a deceptively obvious solution to human error; removing the human. While unthinkable even a few years ago, Artificial Intelligence may present a road to this end. And, with it, plenty of potholes and hazards to trip those with even the purest intentions. The law lumbers forward while technology sprints; a system as envisioned in this article might be inevitable, but early-adopters must appreciate the uncharted waters they navigate. Section II first explores the current state of sexual harassment law, and why it fails the harassed. Section III then examines one possible solution—training Artificial Intelligence to recognize sexual harassment. Section IV then explores the legal implications

---

16. Stefanie K. Johnson, Jessica F. Kirk, & Ksenia Keplinger, *Why We Fail to Report Sexual Harassment*, HARV. BUS. REV. (Oct. 4, 2016), https://hbr.org/2016/10 /why-we-fail-to-report-sexual-harassment [https://perma.cc/939C-W653].

17. For instance, 21st Century Fox maintains a confidential sexual harassment hotline. And yet, despite the abundance of sexual harassment allegations now swirling around Bill O'Reilly, no one ever reported him using the hotline. Sarah Kessler, *Corporate Sexual Harassment Hotlines Don't Work. They're Not Designed to,* QUARTZ AT WORK (May 2, 2017), https://work.qz.com/971112/corporate-sexual-harassment-hotlines-dont-work-theyre-not-designed-to [https://perma.cc/89YV-JWE Q].

18. Barbara Ortutay, *Can Better Policies Prevent Workplace Sexual Harassment?*, AP NEWS (Dec. 8, 2017), https://www.apnews.com/45ff9a0e936e43 c6a9fad9f892b7872e [https://perma.cc/6X62-MLGN] (quote from Dan Eaton). Of course, there are plenty of other solutions. Microsoft recently rescinded forced arbitration of sexual harassment claims, encouraging those who experience harassment to seek relief. Chris Morris, *Microsoft Changes Its Sexual Harassment Policies in the Wake of #MeToo*, FORTUNE (Dec. 19, 2017, 8:33 AM), http://fortune.com/2017/12/19/microsoft-changes-sexual-harassment-policies/ [https://perma.cc/2YDG-UZ2H]. Google uses unconscious bias training to educate employees on recognizing issues within their own behavior. Renzo Costarella, *These Companies Are Battling Sexual Harassment by Teaching Employees to Recognize Unconscious Bias*, ENTREPRENEUR (Dec. 1, 2017), https://www.entrepreneur. com/article/305324 [https:// perma.cc/4BQY-DD5B]. And bystander training remains a popular option for tackling the issues of underreporting. Johnson, Kirk, & Keplinger, *supra* note 16.

of such a system, both in expanding sexual harassment law and its implications for privacy, reputation, and workplace comradery.

II. "The Fault in Our Prongs": Why Current Law Fails the Harassed

> "I just start kissing them. It's like a magnet. Just kiss. I don't even wait. And when you're a star, they let you do it. You can do anything . . . . Grab 'em by the pussy. You can do anything."[19]

Title VII's purpose is—and always was—prophylactic. Congress was not trying to provide victims redressability.[20] Rather, Congress wished to "assure equality of employment opportunities and to *eliminate* those discriminatory practices and devices which have fostered" barriers like those built on race and sex.[21] To achieve that, it "arm[ed] the courts with full equitable powers" to "make persons whole for injuries suffered on account of unlawful employment discrimination."[22] Redressability would, in turn, "influence primary conduct" to "avoid harm."[23] Harassment costs businesses money, and money motivates change.[24] Oddly, slightly misguided jurisprudence has muddled this harm-avoidance principle.

A. Sexual Harassment is Not a Character Flaw

Title VII's language is deceptively clear—"[i]t shall be an unlawful employment practice for an employer . . . to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's . . . sex . . . ."[25] In other words, employers cannot

---

19. *Transcript: Donald Trump's Taped Comments About Women*, N.Y. TIMES (Oct. 8, 2016), https://www.nytimes.com/2016/10/08/us/donald-trump-tape-transcript.html [https://perma.cc/QAC6-LW55] (quote from Donald J. Trump).

20. McDonnell Douglas Corp. v. Green, 411 U.S. 792, 800 (1973).

21. *Id.* (emphasis added).

22. Albemarle Paper Co. v. Moody, 422 U.S. 405, 418 (1975).

23. Faragher v. City of Boca Raton, 524 U.S. 775, 805–06 (1998).

24. *See* Parramore, *supra* note 9.

25. 42 USC § 2000e-2(a)(1) (2018). Conventional wisdom holds Congress added sex discrimination at the eleventh hour as a joke. *See, e.g.*, Diaz v. Pan Am. World Airways, Inc., 442 F.2d 385, 386 (5th Cir. 1971) ("The amendment adding the word 'sex' to 'race, color, religion and national origin' was adopted one day before House

discriminate against someone *because of* their sex. But the first courts to adjudicate sexual harassment under Title VII grappled with this idea that harassment was, itself, a form of discrimination.[26] Instead, Title VII was meant to "strike at the entire spectrum of disparate treatment of men and women resulting from sex stereotypes."[27] Handsy managers, to these early courts, had little to do with sex stereotypes.[28]

Barnes v. Train was the first opinion to consider sexual harassment under Title VII.[29] Plaintiff "refused to engage in a sexual affair with her supervisor,"[30] and was belittled, harassed, and eventually fired as a result.[31] But, to the *Barnes* court, "[r]egardless of how inexcusable the conduct of plaintiff's supervisor might have been, it does not evidence *an arbitrary barrier* to continued employment based on plaintiff's sex."[32] Instead, the harassment was a "controversy underpinned by the subtleties of an inharmonious personal relationship."[33]

Corne v. Bausch & Lomb, Inc. was the first reported case considering sexual harassment under Title VII.[34] Again, the facts outline a clear claim. Plaintiffs Corne and DeVane were clerical workers and "repeatedly subjected to verbal and physical sexual advances from defendant Price[,]" which made work so intolerable

---

passage of the Civil Rights Act. It was added on the floor and engendered little relevant debate."). However, this popular account has been revealed as a "misleading and harmful" reduction of hard-fought rights. Vicki Schultz, *Taking Sex Discrimination Seriously*, 91 DENV. U. L. REV. 995, 1020 (2015).

26. *See* Barnes v. Train, 1974 U.S. Dist. LEXIS 7212, at *3; Corne v. Bausch & Lomb, Inc., 390 F. Supp. 161, 162 (D. Ariz. 1975).

27. Sprogis v. United Air Lines, Inc., 444 F.2d 1194, 1198 (7th Cir. 1971).

28. This approach, while wrong, is more conservative and evidences the Court's general unwillingness to target personal behavior. Private employers were within the purview of the statute—private people were not.

29. *Sexual Harassment and Title VII: The Foundation for the Elimination of Sexual Cooperation as an Employment Condition*, 76 MICH. L. REV. 1007, 1007 n.3 (1978).

30. *Barnes*, 1974 U.S. Dist. LEXIS 7212, at *3. The original opinion was fairly scant, but the appellate decision overturning *Barnes* is thankfully more detailed. Barnes v. Costle, 561 F.2d 983, 984–86 (1977).

31. *Barnes*, 561 F.2d at 984–86.

32. *Barnes*, 1974 U.S. Dist. LEXIS 7212, at *3 (emphasis added).

33. *Id.*

34. Corne v. Bausch & Lomb, Inc., 390 F. Supp. 161 (D. Ariz. 1975). *Barnes* was not widely known until it was overturned in 1977. Instead, *Corne* was the first reported case of "a sexual harassment claim under Title VII," and the first to gain any notoriety. CATHARINE A. MACKINNON, SEXUAL HARASSMENT OF WORKING WOMEN 60 (1979).

that they left.[35] But the court denied them relief, concluding "Price's conduct appears to be nothing more than a personal proclivity, peculiarity[,] or mannerism."[36]

These courts were correct on one count—Title VII is prophylactic. It is meant to "make careers open to talents irrespective of race or sex," rather than simply providing redress for injuries.[37] Lawsuits are a means to that end, rather than the end itself. But when the company "can only be damaged by the very nature of the acts complained of,"[38] Title VII is irrelevant as motivation to correct poor behavior already exists.[39]

It was not until the mid-seventies that district courts began changing their tune. *Williams v. Saxbe* signaled this shift in its groundbreaking finding that a male supervisor's retaliatory acts could "create[] an artificial barrier to employment . . . placed before one gender and not the other[.]"[40] However, the court refused to equate sexual harassment and discrimination as a matter of law; rather, it was up to the fact-finder to determine whether harassment was "a policy or a regulation of the office . . . [or] an isolated personal incident."[41]

---

35. *Corne*, 390 F. Supp. at 162.

36. *Id.* at 163.

37. Tomkins v. Pub. Serv. Elec. & Gas Co., 422 F. Supp. 553, 556 (D.N.J. 1976) *overruled by* Tomkins v. Pub. Serv. Elec. & Gas Co., 568 F.2d 1044 (3d Cir. 1977). This lower court's opinion neatly summarizes the general view of Title VII under early jurisprudence:

> Title VII was enacted in order to remove those artificial barriers to full employment which are based upon unjust and long-encrusted prejudice. Its aim is to make careers open to talents irrespective of race or sex. It is not intended to provide a federal tort remedy for what amounts to physical attack motivated by sexual desire on the part of a supervisor and which happened to occur in a corporate corridor rather than a back alley.

Tomkins, 422 F. Supp. at 556.

38. *Corne*, 390 F. Supp. at 163.

39. The court required "discriminatory conduct . . . ar[i]se out of company policies" where "[t]here was apparently some advantage to, or gain by, the employer from such discriminatory practices." *Id.* Interestingly, the *Corne* court's concerns mirror many voiced about the #MeToo Movement: namely that "an outgrowth of holding such activity to be actionable under Title VII would be a potential federal lawsuit every time any employee made amorous or sexually oriented advances toward another. The only sure way an employer could avoid such charges would be to have employees who were asexual." *Id.* at 163–64.

40. Williams v. Saxbe, 413 F. Supp. 654, 657 (D.D.C. 1976).

41. *Id.* at 660. A similar decision appeared in the Fourth Circuit in Garber v. Saxon Bus. Prod's, wherein the appellate court found the complaint, liberally construed, could show "an employer policy or acquiescence in a practice of compelling female employees to submit to the sexual advances of their male supervisors . . . ." 552 F.2d 1032, 1032 (4th Cir. 1977).

The Third Circuit in *Tomkins* distilled the key difference between earlier cases and recent decisions like *Saxbe*, by focusing on adverse action.[42] Harassment moves away from personal conflict and towards discrimination once "direct employment consequences flow[] from the advances[.]"[43] And, in the court's opinion, foreshadowed the eventual legal framework under which Title VII operates.[44] Once (1) "a supervisor," (2) with "actual or constructive knowledge of the employer," (3) "makes sexual advances or demands toward a subordinate[,]" and (4) "conditions that employee's job status . . . on a favorable response to those advances[,]" Title VII is violated as a matter of law unless (5) "the employer . . . take[s] prompt and appropriate remedial action after acquiring such knowledge."[45]

Confusion, rather than clarity, unfortunately reigned post-*Tomkins*. Many courts focused unnecessarily on the harasser's supervisory authority; unless the harasser expressly conditioned continued employment on acquiescence and had the authority to follow through on his threat, the harassment claim failed.[46] Others happily permitted a Title VII claim, even when the harassment came from a co-worker, as long as the employer knew about it and did not step in.[47] Inconsistent application led the EEOC to issue new guidelines in 1980 declaring both *quid pro quo* and hostile environment sexual harassment claims are actionable sex discrimination.[48]

---

42. *Tomkins*, 568 F.2d at 1046.

43. *Id.* at 1048.

44. *Id.*

45. *Id.* at 1048–49.

46. *See, e.g.*, Fisher v. Flynn, 598 F.2d 663, 666 (1st Cir. 1979) (dismissing because the plaintiff, a professor fired after refusing "to accede to the romantic overtures" of her academic department's chairman, could not state a Title VII claim as she failed to allege the chairman had any "authority to terminate her employment or effectively recommend the same"); Clark v. World Airways, 1980 U.S. Dist. LEXIS 15379, at *8–9 (D.D.C. Oct. 24, 1980) (noting "explicit sexual advances" made by a company's president towards an employee were not actionable because the president "did not at any time relate submission to his advances to any of plaintiff's employment prospects"); Smith v. Rust Eng'g Co., 1978 U.S. Dist. LEXIS 14718, at *3 (N.D. Ala. Oct. 25, 1978) (holding that "acquiescence in the sexual advances" of a co-worker was not "impliedly or expressly required as a condition of employment" and thus Plaintiff could not base her Title VII claim on that harassment).

47. *See, e.g.*, Kyriazi v. W. Elec. Co., 461 F. Supp. 894, 924–36 (D.N.J. 1978) (employer violated Title VII when it learned plaintiff's co-workers "created an obscene cartoon" to "humiliate her, AS a woman," but failed to take remedial action).

48. 29 C.F.R. § 1604.11 (1981).

The EEOC identified that all of these courts were effectively enshrining principles of "respondeat superior . . . that an employer is responsible for the acts of its supervisors and agents" in a sexual harassment context.[49] Obviously, that meant harassment by an employer's direct agent imputed liability to the company. But it also meant the company could be liable for harassment by non-supervisory employees, as long as the company knew about it and did nothing.[50]

These guidelines lay the root of all modern sexual harassment litigation, and their dual-concerns evidences the inherent conflict this paper seeks to address. The guidelines' "major thrust" is "[p]revention is the best tool for the elimination of sexual harassment."[51] Litigation was the stick; prevention the carrot. Once the employer "knows of acts of sexual harassment . . . and rectifies the actual results of those actions," workers could no longer wield the stick.[52]

B. The Prongs Cometh

As noted by the Supreme Court in *Meritor Sav. Bank, FSB v. Vinson,* "[s]ince the Guidelines were issued, courts have uniformly held" an employer violates Title VII either when a supervisor sexually harasses a subordinate, or the employer knowingly permits "a hostile or abusive work environment" based on sex.[53] Rectifying sexual harassment is key to removing "arbitrary barrier[s] to sexual equality at the workplace."[54] After all, the harm of sexual harassment extends beyond economic injury as "run[ning] a gauntlet of sexual abuse in return for the privilege of being allowed to work and make a living can be as demeaning and disconcerting as the harshest of racial epithets."[55]

---

49. *Sex Discrimination in the Workplace, 1981: Hearing Before the S. Comm. on Labor and Human Res.* 97th. Cong. 393 (1981) [hereinafter *Hearing*] (statement of J. Clay Smith Jr., Acting Chairman, Equal Emp't Opportunity Comm'n).

50. "However, where both knowledge and control do exist on the part of the employer, there is an obligation under [T]itle VII for the employer to maintain an atmosphere that is free of sexual harassment, so that members of one sex are not required to work under different and less advantageous terms and conditions of employment than members of the other sex." *Id.* at 340.

51. *Id.* at 341.

52. As noted by the EEOC, "the Commission would not sue for a remedy which has already been granted." *Id.* at 340.

53. Meritor Sav. Bank, FSB v. Vinson, 477 U.S. 57, 66 (1986).

54. *Id.* at 67 (quoting Henson v. Dundee, 682 F.2d 897, 902 (1982)).

55. *Id.*

Encouraging employers who willingly dismantle that gauntlet by adopting prophylactic measures, then, is a legitimate goal. And the legally recognized affirmative defenses to sexual harassment are meant to push employers down that road. Originally, employers had one defense—remedial action.[56] Once an employer rectifies a problem, there is no reason to "sue for a remedy which has already been granted."[57] This defense still exists when a plaintiff's co-workers commit the harassment.[58]

However, when the alleged harasser is an employee's supervisor, courts disagreed on whether absolute liability attached under the theory of *respondeat superior*, or if a defense was still available.[59] In *Meritor,* the Supreme Court denounced any absolute rule, and instead instructed lower courts "to look to agency principles for guidance in this area."[60] This instruction led to chaotic results until,[61] in the late nineties, the Supreme Court created a new affirmative defense.[62] First, it recognized absolute liability attached when a supervisor sexually harasses an employee, and subsequently takes adverse action against her.[63] But, when there is

---

56. *Hearing, supra* note 49, at 351.

57. *Id.*

58. 29 C.F.R. § 1604.11(d) (1986) ("With respect to conduct between fellow employees, an employer is responsible for acts of sexual harassment in the workplace where the employer (or its agents or supervisory employees) knows or should have known of the conduct, unless it can show that it took immediate and appropriate corrective action.").

59. Faragher v. City of Boca Raton, 524 U.S. 775, 793 (1998).

60. *Meritor,* 477 U.S. at 72 (1986).

61. The instruction led to three approaches to vicarious liability rooted in agency principles. An employer could be vicariously liable when (1) "supervisors were acting within the scope of their employment when they engaged in the harassing conduct"; (2) "they were significantly aided by the agency relationship in committing the harassment"; or (3) the employer had actual or constructive "knowledge of the harassment." *Faragher*, 524 U.S. at 775, 793. Some circuits would refuse to attach vicarious liability to conduct "motivated solely by individual desires and serves no purpose of the employer" while others focused more heavily on foreseeability and whether the "sexual misconduct arose from or was in some way related to the employee's essential duties." *Id.* at 776, 796.

62. This defense was explained in companion cases Burlington Indus. v. Ellerth, 524 U.S. 742 (1998), and *Faragher*, 524 U.S. at 775.

63. *Faragher*, 524 U.S. at 807; *Ellerth*, 524 U.S. at 765. Interestingly, even post-*Ellerth*, not every state patterned their sexual harassment jurisprudence on Title VII. Until recently, New Jersey attached strict liability to supervisor harassment even without an adverse employment action. Aguas v. State, 107 A.3d 1250, 1278 (N.J. 2015) (Albin, J., dissenting) ("Until today, the *Ellerth/Faragher* affirmative defense standard was foreign to our . . . jurisprudence . . . [which] makes no such distinction between tangible employment actions and sexual harassment that may cause physical or psychological harm to the employee.").

no adverse action, an employer may avoid liability by showing (1) they "exercised reasonable care to prevent and correct promptly any sexually harassing behavior"; and (2) "that the plaintiff employee unreasonably failed to take advantage of any preventive or corrective opportunities provided by the employer."[64]

The two requirements of an *Ellerth/Farragher* defense creates different obligations for each party; for simplicity's sake, this paper refers to each obligation as prongs one through three.[65] The first prong obliges employers use "reasonable care to prevent . . . sexually harassing behavior."[66] Prong two mandates employers minimize harm once it occurs.[67] And, if these prongs are met, prong three's burden shifts to plaintiff to prove they either took advantage of preventative measures, or their failure to do so was reasonable.[68]

C. Shouldering Your Burden

Courts have recognized sexual harassment as a form of sex discrimination for nearly thirty years. And yet, as the #MeToo movement shows, harassment remains one of the largest arbitrary barriers to gender equality in the workplace. So what went wrong?

Some analysts blame the defense itself. They derisively dub it the "file cabinet defense," requiring employers do little beyond "promulgat[ing] an anti-harassment policy that specifically addresses sexual harassment and a grievance procedure that allows an employee to bypass a harassing supervisor."[69] As long as you

---

64. *Faragher*, 524 U.S. at 807; *Ellerth*, 524 U.S. at 765.

65. Traditionally, there are only two prongs of the *Ellerth/Faragher* defense: (1) The employer must use reasonable care to prevent and correct harassing behavior; and (2) the employee must take advantage of any preventive or corrective opportunities. *See Faragher*, 524 U.S. at 807. However, one of the biggest failings of sexual harassment law, in the author's opinion, if this tendency to group prevention and correction together. Prevention must remain the end-goal; correction simply fills the gaps.

66. *Faragher*, 524 U.S. at 807.

67. *Id.*

68. *Faragher*, 524 U.S. at 807–08.

69. Anne Lawton, *Operating in an Empirical Vacuum: The Ellerth and Faragher Affirmative Defense,* 13 COLUM. J. GENDER & L. 197, 210 (2004) ("My concern was well-founded. An examination of 200 federal court cases addressing the new *Ellerth* and *Faragher* affirmative defense reveals that many federal courts have interpreted the Court's decisions in *Ellerth* and *Faragher* to require little more than what the Court in *Meritor* commanded: promulgate an anti-harassment policy that specifically addresses sexual harassment and a grievance procedure that allows an employee to bypass a harassing supervisor. As a practical matter, the Court's decisions in *Ellerth* and *Faragher* did little to change employer incentives to reduce the incidence of

have a file cabinet for these claims, the employee bears the burden of acting. A fairer assessment views this as a "notice" requirement,[70] as exists in some premises liability cases.[71] A landlord has a duty to monitor for hazardous conditions but, as long as that duty is satisfied, she is not liable for injury caused by unknown conditions.

But requiring notice is futile when employees refuse to report. There is a deep fear of retaliation.[72] Seventy-six percent of those who report harassment experience retaliation, which certainly hinders others from reporting.[73] Indeed, as one analyst notes, "if the vast majority of harassment victims do not report harassment, then the reasonable response is *not* to report harassment."[74] The reporting requirement stems from tort-principles of mitigation; a victim is only blameless as long as damage is unavoidable.[75] Harm-avoidance should certainly remain part of a court's analysis, but current laws unnecessarily constrain harm-avoidance to formal reporting.[76] And, as a result, many stories go unheard.

---

sexual harassment by supervisors in the workplace.").

70. The Ninth Circuit nicely summarized the notice requirement as follows:
   Notice of the sexually harassing conduct triggers an employer's duty to take prompt corrective action that is 'reasonably calculated to end the harassment.' This obligation actually has two parts. The first consists of the temporary steps the employer takes to deal with the situation while it determines whether the complaint is justified. The second consists of the permanent remedial steps the employer takes once it has completed its investigation.
Swenson v. Potter, 271 F.3d 1184, 1192 (9th Cir. 2001) (internal citations omitted).

71. The Americans with Disabilities Act provides another analogous legal mechanism. Designed to "eliminat[e] . . . discrimination against individuals with disabilities;" the ADA requires an employer provide "reasonable accommodations" to help a disabled but otherwise qualified individual to do their job. 42 U.S.C. § 12101(b) (2018). However, "[a]n employee has the initial duty to inform the employer of a disability before ADA liability may be triggered for failure to provide accommodations—a duty dictated by common sense lest a disabled employee keep his disability a secret and sue later for failure to accommodate." Beck v. Univ. of Wis. Bd. of Regents, 75 F.3d 1130, 1134 (7th Cir. 1996).

72. Johnson, Kirk, & Keplinger, *supra* note 16.

73. *See id.*; *See also* Cortina & Magley, *supra* note 15, at 255 (finding that only 34% of those who reported harassment did not experience any form of retaliation).

74. Lawton, *supra* note 69, at 209.

75. Burlington Indus. v. Ellerth, 524 U.S. 742, 764 (1998) ("Title VII borrows from tort law the avoidable consequences doctrine, . . . and the considerations which animate that doctrine would also support the limitation of employer liability in certain circumstances.").

76. Margaret Johnson explores this phenomenon in depth, noting "the acts taken to 'avoid harm' from sexual harassment are more diverse than filing a formal complaint of sexual harassment." Margaret E. Johnson, *"Avoiding Harm Otherwise": Reframing Women Employees' Responses to the Harms of Sexual Harassment*, 80 TEMP. L. REV. 743, 771 (2007). Johnson suggests lawmakers and judges look beyond formal reporting mechanisms and recognize how "more complete stories can be told

Further, harassment is not a one-size-fits-all problem. Not all misconduct is actionable, nor do all employees perceive harassment the same way. Disagreement abounds, even within the #MeToo movement.[77] No one really knows when, exactly, a compliment crosses the line into harassment.[78] There are simply too many subjective elements.

Scholars have proposed different solutions to these two problems. Some suggest a new reasonableness standard, permitting juries to determine the reasonableness of an employee's failure to report.[79] Others argue the standard should remain with the employer, requiring it prove it acted reasonably in trying to prevent harassment.[80] But these solutions overlook an interesting quality of sexual harassment law. The interests of the employer and employee are not diametrically opposed. Sexual harassment harms both.[81] A cooperative system that recognizes allies (the harassed and the employer) and the problem (the harasser) could solve underreporting. Retaliation and confusion shrink against a company's vast resources and honest support. Perhaps, then, it is reasonable to expect the employer do more in helping employees avoid harm.

---

and made available for determinations pursuant to the liability framework for supervisor sexual harassment." *Id.* at 806.

77. Fiona Chen, *Why the Aziz Ansari Story and Discussions of Grey Areas Are Central to the #MeToo Movement*, TECH (Jan. 25, 2018), https://thetech.com/2018/01/25/me-too-aziz-ansari [perma.cc/B8FM-69YV].

78. Akshita Jha & Radhiki Mamidi, *When Does a Compliment Become Sexist? Analysis and Classification of Ambivalent Sexism Using Twitter Data*, PROCEEDINGS OF THE SECOND WORKSHOP ON NAT. LANGUAGE PROCESSING AND COMPUTATIONAL SOC. SCI. 7 (2017) ("Previous works on computationally detecting sexism present online are restricted to identifying the hostile form. Our objective is to investigate the less pronounced form of sexism demonstrated online.").

79. Heather S. Murr, *The Continuing Expansive Pressure to Hold Employers Strictly Liable for Supervisory Sexual Extortion: An Alternative Approach Based on Reasonableness*, 39 U.C. DAVIS L. REV. 529, 635 (2006).

80. Robert R. Graham III, *Two Wrongs Do Not Make a Right: The Need to Revisit the Ellerth/Faragher Affirmative Defense*, 30 NOTRE DAME J.L. ETHICS & PUB. POL'Y 423, 446 (2016) ("An employee's reporting of harassing behavior creates a rebuttable presumption of liability that shifts the burden to the defendant employer to prove that it acted reasonably in its efforts to prevent and remedy the harassment at issue.").

81. *See* Parramore, *supra* note 9.

III. "Do Androids Dream of Electric Sheep?": Teaching the Robot

> "Did you know that my love is a liquid? I could talk to me for years. I can't speak to you at all. Did you know that friends come in boxes?"[82]

At a 1956 conference at Dartmouth College, R.J. Solomonoff posited "[t]he finding of 'words' that have a high and useful correlation with other 'words' has been, by far, the most important task of science."[83] When someone walks outside and sees dark clouds, their mind will predict the likelihood of rain because it is hardwired to recognize patterns.[84] A result—rain—is partly predictable based on an element—clouds. Add more elements—temperature, time of day, geography—and a more accurate prediction follows. According to Solomonoff, programmers could theoretically create an algorithm to examine elements, identify corresponding results, and then issue predictions.[85]

A. Making Intelligence Artificial: Algorithmic Learning

Artificial Intelligence is an incredibly complex and eclectic field. Solomonoff was one of ten academics at Dartmouth, and a schism immediately appeared between the "rigidly logical" but "immediately applicable" programs, and probabilistic based programs.[86] Systems are, in fact, still categorized as driven by either logic or probability.[87]

---

82. TUBEWAY ARMY, *My Love is a Liquid*, *on* TUBEWAY ARMY (Beggars Banquet Records 1978).

83. Ray Solomonoff, An Inductive Inference Machine, IRE Convention Record, Section on Information Theory, Part 2, at 2 (1957).

84. Or, put another way, "if we observe 'cloudy weather,' we expect 'rainy weather' in the immediate space-time neighborhood[—]and vice-versa." *Id.*

85. "To make predictions, the machine examines the interrogation square of the q element, finds a p n-gram that fits the problem, and makes the prediction that corresponds to that p n-gram." *Id.* at 14.

86. Grace Solomonoff, *Ray Solomonoff and the New Probability*, *in* ALGORITHMIC PROBABILITY AND FRIENDS, BAYESIAN PREDICTION AND ARTIFICIAL INTELLIGENCE: PAPERS FROM THE RAY SOLOMONOFF 85TH MEMORIAL CONFERENCE, MELBOURNE, VIC, AUSTRALIA, NOV./DEC. 2011, 37, 42 (David L. Dowe ed., 2013) (explaining that logic-based programs are defined by much simpler "if-then" statements; if option a, then result y, but if option b, result x).

87. *See* Karthik Guruswamy, *Data Science: Machine Learning Vs. Rules Based Systems*, FORBES (Dec. 15, 2015), https://www.forbes.com/sites/teradata/2015/12/15/data-science-machine-learning-vs-rules-based-systems/#2f5b61b2119a [https://perma.cc/5NX6-DSZQ]. However, many programs, like IBM's Watson, actually use

The biggest drawback to purely logical programming is the difficulty in hard-coding knowledge.[88] Our world is inherently complex, and accurately distilling complexity into logical rules is almost impossible.[89] Machines needed to acquire their own knowledge "by extracting patterns from raw data," much like humans do.[90] Machine learning is now the most cutting-edge form of AI as it takes away the human operator, and theoretically permits a program to simply learn on its own.[91] A machine with enough computing power can break down even complicated collections of information and, from those, extract patterns.[92]

Pattern extraction is AI's most interesting feature, particularly in preventing sexual harassment. AI is exceedingly good at mapping out extant and predicted relationships based on these patterns.[93] Police departments have dubbed this "predictive policing," as their software pores over data to predict potential criminals, victims, and co-offending networks.[94] Credit card monitoring software pores through billions of transactions for

---

both. John Rennie, *How IBM's Watson Computer Excels at Jeopardy*, PLOS (Feb. 14, 2011), http://www.cs.cornell.edu/courses/cs6700/2013sp/readings/01-a-Watson-Short .pdf [https://perma.cc/K7FJ-TQZ9].

88. *See* IAN GOODFELLOW ET AL., DEEP LEARNING 2 (2016).

89. *Id.* at 2–3.

90. *Id.*

91. *Id.* at 1–3.

92. Humans already do this subconsciously. Imagine, for instance, a cat. Your mind can generally predict how a cat looks based on your past interactions with the animal. But AI programs could not mimic this behavior without first being told what a cat looked like. Then, in 2012, researchers behind Google Brain discovered their AI was able to successfully compile a probabilistic image of a cat purely through pattern recognition after reviewing ten million random video thumbnails. Google's team never told the machine that a certain object was a cat. Rather, the machine recognized the pattern as significant within the majority of the random thumbnails, and assigned its conception of a cat value. *See* John Markoff, *How Many Computers to Identify a Cat? 16,000*, N.Y. TIMES (June 25, 2012), https://www.nytimes.com/201 2/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.ht ml [https://perma.cc/W65S-RNHL].

93. *See China: Police 'Big Data' Systems Violate Privacy, Target Dissent*, HUMAN RIGHTS WATCH (Nov. 19, 2017), https://www.hrw.org/news/2017/11/19/china-police- big-data-systems-violate-privacy-target-dissent [https://perma.cc/7WDW-465G].

94. Social media is a treasure trove of data for predicting would-be criminals and mapping their relationships. *See* MOHAMMAD A. TAYEBI & UWE GLÄSSER, SOCIAL NETWORK ANALYSIS IN PREDICTIVE POLICING: CONCEPTS, MODELS AND METHODS 2 (2016). This practice, of course, comes with its own set of problems, including the implicit bias inherent in any community-based prediction software. Randy Rieland, *Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?*, SMITHSONIAN MAG. (Mar. 5, 2018), https://www.smithsonianmag.com/innovation/artificial-intellig ence-is-now-used-predict-crime-is-it-biased-180968337/ [https://perma.cc/TU5B-S9L M].

patterns to indicate someone is a thief or terrorist.[95] The Chinese Government has used machine-learning programs for nearly everything, from planning city routes to rooting out political dissidents.[96]

Given that a lot of companies are using AI to replace core functions once left to Human Resources,[97] a program that similarly predicts sexual harassment could be a godsend. Sexual harassment software already has a place in corporate America. Programs analyze the content of e-mails to warn users away from sending potentially harassing messages.[98] Advanced Discovery's new "Riskcovery" actually uses predefined patterns of sexually harassing behavior to flag potential harassment.[99] But the problem with these programs, at least from a preventative perspective, is they still require harassment to occur, or almost occur, before sending up a flag. If AI can truly vindicate Title VII's prophylactic purpose, it will need to kick in earlier in the process.

B. *Can You Teach AI to be Harassed?*

Sexual harassment is predictable. Socio-cultural studies prove researchers can analyze social situations, personalities, and contextual clues to determine when sexual harassment might

---

95. *See* Robert J. O'Harrow, Jr., *No Place to Hide*, 24 J. MARSHALL J. COMPUTER & INFO. L. 35, 45 (2005).

96. *See* Liza Lin & Josh Chin, *China's Tech Giants Have a Second Job: Helping Beijing Spy on Its People*, WALL ST. J. (Nov. 30, 2017), https://www.wsj.com/articles/c hinas-tech-giants-have-a-second-job-helping-the-government-see-everything-15120 56284 [https://perma.cc/T2CY-439A]; Sam Schechner, Douglas MacMillan, and Liza Lin, *U.S. and Chinese Companies Race to Dominate AI*, WALL ST. J. (Jan. 18, 2018),https://www.wsj.com/articles/why-u-s-companies-may-lose-the-ai-race-151628 067 [https://perma.cc/T29Y-7UNX].

97. *See* Jeanne Meister, *The Future Of Work: The Intersection of Artificial Intelligence and Human Resources*, FORBES (Mar. 1, 2017), https://www.forbes.co m/sites/jeannemeister/2017/03/01/the-future-of-work-the-intersection-of-artificial-intelligence-and-human-resources/2/#49c2b98c67ee [https://perma.cc/XCA4-WRBP]; Ted Greenwald, *How AI Is Transforming the Workplace*, WALL ST. J. (Mar. 10, 2017), https://www.wsj.com/articles/how-ai-is-transforming-the-workplace-1489371060 [https://perma.cc/3CZG-CRJP]; Sam Schechner, *Meet Your New Boss: An Algorithm*, WALL ST. J. (Dec. 10, 2017), https://www.wsj.com/articles/meet-your-new-boss-an-algorithm-1512910800 [https://perma.cc/6FJY-GUSP].

98. *See* Joanna Goodman, *Legal Technology: Better Together*, LAW SOC'Y GAZETTE, July 11, 2016, https://www.lawgazette.co.uk/features/legal-technology-better-together/5056456.article [https://perma.cc/KB4U-KCLS].

99. *See Riskcovery & Compliance Risk Assessment*, ADVANCED DISCOVERY, https://www.advanceddiscovery.com/what-we-do-for-you/risk-management-complia nce/riskcovery-compliance-risk-assessment/ [https://perma.cc/6NNU-RDD6].

occur.[100] The trouble is reducing these highly malleable, subjective experiences into language understood by the program. Just as Google Brain had to learn the very concept of a cat,[101] so too would an effective AI need to understand the concept of harassment. However, few would disagree on the very basic ingredients necessary in a cat.[102] Many disagree as to what constitutes harassment.[103]

Subjective experience long-confounded AI researchers as rudimentary AIs required binary definitions. "Honesty is good, dishonesty is bad" was a very simple moral game that effectively ignores subjective experience and context. However, with the advent of neural networks, there can be ranges with small upticks between each judgment. And it is here that researchers are having luck teaching AI about human culture.[104] AIs have, in fact, been taught to accommodate human culture before determining if an act

---

100. *See* John B. Pryor et al., *A Social Psychological Model for Predicting Sexual Harassment*, 51 J. Soc. Issues 69, 70 (1995).

101. *See* Markoff, *supra* note 92.

102. Even if you prefer dogs to cats, you and I probably agree on what a cat looks like.

103. *See* Wendy L. Patrick, *Sexual Harassment Is in the Eye of the Beholder*, Psychol. Today (Mar. 15, 2017), https://www.psychologytoday.com/us/blog/why-bad-looks-good/201703/sexual-harassment-is-in-the-eye-the-beholder [https:// perm a.cc/PQV3-6H74].

104. Researchers were able to teach a machine to mimic cultural preferences in negotiation using games. The goal was to move across a colored board to a 'goal' square. But moving required surrendering a chip of whatever color adjacent square you wanted to go to. There were three phases: negotiation, transfer, and movement. During the negotiation phase, a party could ask the other to surrender chips of a certain color. If rejected, the other side got to ask. During transfer, you could transfer however many chips you agreed to, or a subset of that agreement (including none), allowing you to break the promise. Finally, during the movement phase, you could move one step if you had the chips. After that, the positions alternated and the next round would begin. The game ended when a party either didn't move for three rounds or reached its goal. The game was then scored by how close each party got to its goal. *See* Galit Haim et al., *A Cultural Sensitive Agent for Human-Computer Negotiation*, *in* Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems 451, 452 (June 2012), http://www.eecs.harvard. edu/~gal/Papers/pal.pdf [https://perma.cc/BA7T-9KUQ].

is morally wrong.[105] They can simulate guilt,[106] flirt,[107] and predict potential suicide risks.[108] An AI can even recognize beauty and music, as long as those components have an assignable value.[109]

Title VII is not a civility code,[110] which is why a plaintiff cannot successfully sue over one dirty joke. But the "objective severity of harassment should be judged from the perspective of a reasonable person in the plaintiff's position, considering 'all the circumstances.'"[111] That standard adds a subjective element. The first step, then, in designing an AI program for sexual harassment is teaching it to accommodate subjective experience. Having new hires play a game with the machine could allow the program to assign a 'sensitivity profile' for each employee.[112]

C. Designing "An Eye"[113] Program

This game cannot simply spit out dirty jokes or potentially harassing situations and ask new employees to rank their comfort

---

105. *See* Ariel Conn, *Ethics and Creativity in Artificial Intelligence: An Interview with Mark Riedl*, HUFFINGTON POST (June 1, 2017), https://www.huffingtonpost.co m/entry/ethics-and-creativity-in-artificial-intelligence-an_us_593047b4e4b09e93d7 964848 [https://perma.cc/75PY-GRS6]; *Using Stories to Teach Human Values to Artificial Agents*, GA. TECH. (Feb. 12, 2016), http://www.news.gatech.edu/2016/02 /12/using-stories-teach-human-values-artificial-agents [https://perma.cc/52MB-P6 RY].

106. *See* Simon Parkin, *Teaching Robots Right from Wrong*, ECONOMIST:1843 MAG. (June/July 2017), https://www.1843magazine.com/features/teaching-robots-right-from-wrong [https://perma.cc/PGF5-43XR].

107. *See* Moira Weigel, *Flirting with Humanity*, NEW REPUBLIC (May 11, 2016), https://newrepublic.com/article/133034/flirting-humanity [https://perma.cc/6R4W-VJPS].

108. *See* Aili McConnon, *AI Helps Identify People at Risk for Suicide*, WALL ST. J. (Feb. 23, 2018), https://www.wsj.com/articles/ai-helps-identify-people-at-risk-for-suicide-1519400853 [https://perma.cc/MW86-7YPW].

109. AI analyzes images and art deemed popular and professional, identifies common patterns in the art, and then compares new images to what it found. Levi Manovich, *Automating Aesthetics: Artificial Intelligence and Image Culture*, MANOVICH (2017), http://manovich.net/index.php/projects/automating-aesthetics-art ificial-intelligence-and-image-culture [https://perma.cc/8BZS-XL4X]. As the author points out though, relying on popular reaction may lead to cultural homogenization. The same application, and results, could apply to music. George Sims, *Will AI Change the Future of Music?*, DEUTSCHE WELLE (Oct. 19, 2017), http://www.dw.com/en/will-ai-change-the-future-of-music/a-41031406 [https://perm a.cc/C7SE-AVYT].

110. Burlington N. & Santa Fe Ry. Co. v. White, 548 U.S. 53, 68 (2006) ("Title VII, we have said, does not set forth 'a general civility code for the American workplace.'").

111. Oncale v. Sundowner Offshore Servs*.,* 523 U.S. 75, 81 (1998).

112. *See* Haim, *supra* note 104 (explaining that games are an excellent way to help AI programs understand cultural differences).

113. Please forgive the terrible pun, but the author could not resist the Orwellian

level.[114] Rather, designers should pattern it around the very personality traits that indicate sensitivity to harassment. One study actually tracked these personality traits and linked them to sensitivity.[115] Some findings were expected; those with a high sense of morality or justice may be more sensitive to harassment as they keenly perceive lapses in decorum.[116] And people of lower intelligence "tend to be over-sensitive to harassment – both sexual and non-sexual" as they might lack "enough insight to understand the subtleties in human behavior" and, "[a]s a result, they may perceive harassment in certain behaviours when, in fact, the behaviour may mean something else entirely – humour, for example."[117]

Others were surprising. People who were naturally shy or withdrawn tended to be less sensitive to sexual harassment, and more sensitive to other forms of harassment; the authors posited that these individuals see sexual harassment as something avoidable through withdrawing from a situation, while other forms of harassment require immediate action.[118] People who were outgoing and friendly were, in contrast, highly sensitive to harassment; the authors posited such "'free and easy' styles lead others to believe that they are not at all sensitive, and to deal with them more harshly than . . . others."[119]

The game, then, should test a person's propensity for confrontation, problem-solving skills, and notions of justice. Something as rudimentary as 'Cheat,' wherein both sides may cheat or call out the other for cheating, could target these traits.[120] Of

---

reference.

114. This approach would effectively require new hires to endure harassing behavior, albeit from a machine, after being hired.

115. *See* Stephen M. Crow et al., *The Impact of Personality Factors on Sexual and Non-sexual Harassment Sensitivity*, 10 WOMEN MGMT. REV. 9, 10 (1995).

116. *Id.* at 15–16.

117. *Id.* at 15.

118. *Id.* at 16.

119. *Id.* at 15.

120. For those who never played Cheat—also known as "Bullshit," "Bluff," and "I-Doubt-It"—the game is very simple. The aim is for the player to get rid of all of their cards. Each player takes turns placing cards face down starting from Aces and working their way up, but they must state orally what they are purportedly placing in the pile. For example, the first player will say "two Aces" and place two of their cards on the pile. The other players may then call the first player's bluff or the next player may take their turn. If a player's bluff is called successfully, that player has to pick up the pile. But if the player who called you a bluff is wrong, they pick up the pile.

course, personality tests are not all created equal.[121] But the more games the computer plays, the better the sensitivity profile becomes. That gives the program a base from which to judge patterns and relationships. Is a supervisor with low-sensitivity paired with an employee with high-sensitivity? That immediately raises a red flag, allowing a company to monitor their relationship closely. Being low-or-high sensitivity is not actionable, and it must not influence any managerial decisions,[122] but rather it lets human resources know to watch for issues.

The Eye—like any successful AI algorithm—would require massive amounts of data in order to function properly and would have to learn what harassment actually is before predicting its likelihood.[123] Looking for keywords in e-mails is not enough—no one says "I am going to sexually harass you" in an e-mail. Instead, it must identify the behaviors and situations that have led to allegations of harassment. Programmers have trained programs like Riskcovery and BotlerAI to read publicly available cases and extract patterns from those,[124] but that still requires enough behavior to make a case. Permitting the program to review internal HR files and complaints would help it understand what actually leads to low-level complaints. It would then supplement such a database, creating a centralized location for the aggregate scuttlebutt the program distills from corporate chatter.

Taking this a step further, the program could even reach out of network to comb the internet for all publicly available information about a company's employees. It can use an employee's photo to identify social networks and map out relationships with co-workers based on extant connections, photos, conversations, tags, and content interaction. When a Harvey profile interacts with three-month old content posted by a Sally profile,[125] the program re-

---

121. Whitney Martin, *The Problem with Using Personality Tests for Hiring*, HARV. BUS. REV. (Aug. 27, 2014), https://hbr.org/2014/08/the-problem-with-using-personality-tests-for-hiring [https://perma.cc/9JKS-XAU3].

122. A profile that negatively impacts a person's career would easily lay the basis for a lawsuit.

123. Data is why AI has evolved so exponentially over the last five years. People put more information online than ever before, and an AI becomes better at tailoring its predictions when it has more data. Dave Gershgorn, *Five Years Ago, AI Was Struggling to Identify Cats. Now It's Trying to Tackle 5000 Species*, QUARTZ MEDIA (Apr. 11, 2017), https://qz.com/954530/five-years-ago-ai-was-struggling-to-identify-cats-now-its-trying-to-tackle-5000-species/ [https:// perma.cc/72NX-MXWB].

124. BOTLER AI, https://botler.ai/ [https://perma.cc/KLM7-5UDL]; *Riskcovery, supra* note 99.

125. See the story *supra* Section I.

asses the risk of harassment, and potentially spits out a warning to HR.[126] HR then may choose—or, if company policy dictates, will be required—to step in before harassment occurs. The Eye would track this network on a massive scale, monitoring the behavior and relationships of potentially thousands of employees.[127]

However, the building blocks of such a program are rife with legal issues.[128] This program ranks people based on subjectivity to sexual harassment; categorizes them as potential victims and harassers; consolidates mounds of highly sensitive, private information into one central location; and, perhaps most worryingly, potentially punishes people for acts never committed. The true limitation of such a program, then, is probably not technological. It is legal and, as discussed in the next section, those implications raise the most pressing issues concerning the adoption of a program like The Eye.

IV. "The Call of Higher Duty": AI's Implications for Notice, Prevention, and Privacy

> "Blessed is the man who expects nothing, for he shall never be disappointed."[129]

Any AI program meant to predict and prevent sexual harassment—whether it looks like The Eye or something completely different—must exhibit three features. (1) An early warning system permitting intervention prior to an official report; (2) mapping and categorization of an organization's employees; and (3) consumption of massive amounts of data.

A. *Resurrecting the Duty to Prevent*

As discussed above, the *Ellerth/Faragher* defense requires employers use reasonable care to (1) prevent and (2) correct

---

126. Professor Richard Paul kindly referred to this as a "de Haan warning," which is certainly more palatable than "thought crime."

127. Such a victim-focused solution would also permit HR to check in with the potential harassed to ensure they have not perceived any harassment and help build trust between HR and employees.

128. One could probably transplant the basic components of the author's program and put it in a law school exam to test a student's ability to issue spot.

129. Letter from Alexander Pope to William Fortescue (Sept. 23, 1725), *in* THE WORKS OF ALEXANDER POPE: CORRESPONDENCE 104 (comp. by John Wilson Croker, Vol. IX 1886).

"sexually harassing behavior."[130] Plenty of ink has spilled on the latter requirement, but the former—despite Title VII's prophylactic purpose—remains largely unchanged. Artificial intelligence may change that by (1) creating a duty to monitor, and (2) expanding an employer's duty to warn.

i.   Duty to Monitor

Nearly all federal courts have concluded employers satisfy their preventative obligations simply by creating and disseminating anti-harassment policies.[131] Judges prefer "concentrat[ing] on 'simple, quantifiable standards,'" and none exist in sexual harassment law.[132] The EEOC promulgated guidelines for prevention but, beyond requiring that employers develop a reporting system and train employees about sexual harassment, there is little meat to the regulation.[133] So, as long as the policy is not utterly ineffectual or issued in bad faith, reasonable care is exercised simply by disseminating a sexual harassment policy.[134] Despite some recent movement towards examining an anti-harassment program's effectiveness,[135] nothing incentivizes employers to do more than implement training and take complaints.

---

130. *Ellerth*, 524 U.S. at 765; *Faragher*, 524 U.S. at 778.

131. For an excellent survey of these cases, *see* Lawton, *supra* note 69, at 217–23.

132. *Id.* at 213.

133. EEOC Guidelines on Discrimination Because of Sex, 29 C.F.R. § 1604.11(f) (2019) ("An employer should take all steps necessary to prevent sexual harassment from occurring, such as affirmatively raising the subject, expressing strong disapproval, developing appropriate sanctions, informing employees of their right to raise and how to raise the issue of harassment under title VII, and developing methods to sensitize all concerned."). The EEOC's policy guidance also presents few answers, stating an effective program should "include an explicit policy against sexual harassment that is clearly and regularly communicated to employees" and "encourage victims of harassment to come forward . . . ." Policy Guidance on Current Issues of Sexual Harassment, Equal Emp't Opportunity Comm'n (March 19, 1990), https://www.eeoc.gov/policy/docs/currentissues.html [https://perma.cc/9K62-K87W].

134. *See Faragher*, 524 U.S. at 808 (1998); Caridad v. Metro-N. Commuter R.R., 191 F.3d 283, 295 (2d Cir. 1999) (requiring complaint procedures as part of a policy); Barrett v. Applied Radiant Energy Corp., 240 F.3d 262, 266 (4th Cir. 2001) ("Distribution of an anti-harassment policy provides 'compelling proof' that the company exercised reasonable care in preventing and promptly correcting sexual harassment . . . [t]he only way to rebut this proof is to show that the 'employer adopted or administered an anti-harassment policy in bad faith or that the policy was otherwise defective or dysfunctional.'"); Kohler v. Inter-Tel Techs*.,* 244 F.3d 1167, 1180 (9th Cir. 2001).

135. When New Jersey's Supreme Court adopted the *Ellerth/Faragher* defense for its state's anti-discrimination law, it very clearly stated that "an employer that implements an ineffective anti-harassment policy . . . may not assert the affirmative defense." Aguas v. State, 107 A.3d 1250, 1268 (N.J. 2015) (responding to the dissent's

One possible reason is a lack of alternatives. To many, the best an employer can realistically do is implement training. It was, for decades, logistically impossible to monitor every at-work interaction for potential harassment. But as the price point of artificial intelligence drops, the behavior necessary to satisfy reasonable care could change. Tort law regularly adjusts its standards of reasonable care to include the use of accessible technology, even when that technology is not customarily used by businesses.[136]

Judge Learned Hand famously advocated a mathematical approach to adjusting standards in tort law—courts should weigh costs against severity and likelihood of injury.[137] And that may very well be the standard by which a court weighs the reasonableness of avoiding an AI-based system. Once costs are attainable and outweighed by injury, the standard could reasonably shift. A duty to monitor may become the clear standard by which fact-finders judge a preventative measure's effectiveness.

This conclusion would, in fact, align the standard more closely with agency law. The Second Restatement subjects principals to liability for failure to investigate and warn an agent of

---

opinion that the defense allows employers to hide behind paper anti-discrimination policies). Subsequent decisions scrutinized effectiveness a bit more closely than is typically seen in federal cases. *See* Dunkley v. S. Coraluzzo Petroleum Transporters, 118 A.3d 355, 362 (N.J. Super. Ct. App. Div. 2015) (finding that, while the employer did not seem to monitor the effectiveness of its policies, the fact that managers were proactive and "initiated contact with plaintiff before he uttered a complaint" showed the policy was not mere lip service); Smith v. Hutchinson Plumbing Heating Cooling, No. L-0992-12, 2015 WL 853040, at *7 (N.J. Super. Ct. App. Div. Mar. 2, 2015) ("[D]efendant's proof of lack of training on the policy put its effectiveness in issue and precluded summary judgment to defendant on the basis of the affirmative defense.").

136. *See, e.g.*, Tex. & Pac. Ry. Co. v. Behymer, 189 U.S. 468, 470 (1903) ("What usually is done may be evidence of what ought to be done, but what ought to be done is fixed by a standard of reasonable prudence, whether it usually is complied with or not."); The T. J. Hooper, 60 F.2d 737, 739–40 (2d Cir. 1932) (holding tugboat operator liable for loss of cargo because they did not fit the boat with radio equipment that, while rarely used in the industry, would have given the boat notice of an impending storm); Me. Yankee Atomic Power Co. v. NLRB, 624 F.2d 347, 349 (1st Cir. 1980) ("The 'gravity of the resulting injury' that would occur in the almost unthinkable event of a serious accident requires the owner of a nuclear facility, acting through its employees, to exercise extraordinary vigilance at all times."); Donnell v. Cal. W. Sch. of Law, 200 Cal. App. 3d 715, 725 (Cal. Ct. App. 1988) ("Depending on location, building configurations and technology, Cal Western may have the potential to influence or affect the condition of adjoining property or property at various distances from its own property. However, the existing legal standard does not require Cal Western to seek to exert such influence or effect wherever it may have such potential.").

137. United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947).

unreasonable risks at the worksite.[138] The Occupational Safety and Health Administration (OSHA) embodies this standard by imposing a general duty to monitor worksites and prevent hazardous conditions from developing.[139] An AI program that makes monitoring for sexual harassment as easy as inspecting sites for hazardous conditions could prompt either the court or the EEOC to adopt a similar duty to monitor.[140]

ii. Expanding a Duty to Warn

An interesting side effect to the general blending of the employer's dual-obligations to prevent and correct sexual harassment is the burgeoning appreciation of a "duty to warn." The Restatement of Agency Law has noted a master has duty to warn their servant of imminent harm since 1958.[141] It is not surprising, then, that some courts recognize an affirmative duty to warn about prior sexual harassment. In *Paroline v. Unisys Corp.*, for example, the Fourth Circuit explored the different corrective measures a company could adopt to prevent a known harasser from harassing again.[142] Normally, employers simply fire those who sexually harass

---

138. RESTATEMENT (SECOND) OF AGENCY § 471 (1958).

139. 29 U.S.C. § 654(a)(1) (2018); Metzler v. Arcadian Corp., No. 96-60126, 1997 U.S. App. LEXIS 12693, at *12 (5th Cir. Apr. 28, 1997) ("[F]ocus is on an employer's duty to prevent hazardous conditions from developing in the employment itself or the physical workplace."); Fabi Constr. Co. v. Sec'y of Labor, 508 F.3d 1077, 1081 (D. C. Cir. 2007) ("To establish a violation of the General Duty Clause, the Secretary must establish that: (1) an activity or condition in the employer's workplace presented a hazard to an employee, (2) either the employer or the industry recognized the condition or activity as a hazard, (3) the hazard was likely to or actually caused death or serious physical harm, and (4) a feasible means to eliminate or materially reduce the hazard existed.").

140. OSHA and Title VII are more alike than one may think. Both have a preventative purpose. *Compare* 29 U.S.C. § 651 (2018) ("Congress declares it to be its purpose and policy . . . to assure so far as possible every working man and woman in the Nation safe and healthful working conditions"), *with* Faragher v. City of Boca Raton, 524 U.S. 775, 806 (1998) (noting Title VII exists "to avoid harm."); *see also* MacKinnon *supra* note 34, at 159 (laying the basis for finding sexual harassment as a form of sex discrimination and writing that sexual harassment could actually fall within the jurisdiction of OSHA).

141. Restat 2d of Agency, § 512 (1958) ("If a servant, while acting within the scope of his employment, comes into a position of imminent danger of serious harm and this is known to the master or to a person who has duties of management, the master is subject to liability for a failure by himself or by such person to exercise reasonable care to avert the threatened harm."); *see also* Newman v. Redstone, 354 Mass. 379, 382 (1968) (noting that, while older jurisprudence refused to recognize an affirmative duty to warn, "[t]he trend of modern authority . . . is to the contrary"); Paroline v. Unisys Corp., 879 F.2d 100, 112 (4th Cir. 1989) (recognizing a potential duty to warn).

142. Paroline, 879 F.2d at 107 (noting that under certain circumstances, the court

others.[143] But if they do not, the *Paroline* court held past behavior effectively creates constructive notice for future harassment.[144] Liability could thus attach if a prior-harasser is retained and harasses again, but the employer failed to warn the new victims.[145] Both the Second and Tenth circuits adopted similar versions of the rule, similarly reasoning that prior incidents effectively create constructive notice of future behavior.[146]

 This standard presents the same problem as before—it requires harassment of *someone* before triggering a duty. But it recognizes an employer's preventative obligations, and more importantly lays a foundation for a duty to warn. An artificial intelligence program that monitors employee behavior to predict sexual harassment would create the same constructive notice as in

---

may "impute liability, under certain circumstances, to an employer who failed to take steps to try to *prevent* sexual harassment of the plaintiff.").

 143. The *Paroline* court actually tied the plaintiff's theory to one of negligent retention. *Id.* at 112.

 144. *Id.* at 107 ("An employer's knowledge that a male worker has previously harassed female employees other than the plaintiff will often prove highly relevant in deciding whether the employer should have anticipated that the plaintiff too would become a victim of the male employee's harassing conduct.").

 145. *Id.* at 111 ("Paroline's assertion that the company's failure to warn caused her injury, as well as the fact that Moore met Paroline at work, might be grounds for holding that the injury arose out of her employment."). Although the court did not expressly recognize a duty to warn, subsequent decisions clarified what the Fourth Circuit dubbed "the *Paroline* failure-to-warn theory." *See* Foster v. Univ. of Md.-Eastern Shore, 787 F.3d 243, 255–56 (4th Cir. 2015) (stating that the rule "articulated in *Paroline* remains good law in this Circuit"). That court further explained that an employer has "an affirmative duty to prevent sexual harassment, and will be liable if they 'anticipated or reasonably should have anticipated' that a particular employee would sexual harass a particular coworker and yet 'failed to take action reasonably calculated to *prevent* such harassment.'" *Id.* at 255. However, "for purposes of the *Paroline* failure-to-warn theory, an employer may reasonably rely upon the findings of a state civil rights agency in determining whether an employee poses a risk of creating a hostile work environment"; thus, if a state agency found no evidence of wrongdoing, there is no affirmative obligation to warn. *Id.* at 256.

 146. *Compare* Hirase-Doi v. U.S. W. Commc'ns Inc., 61 F.3d 777, 783 (10th Cir. 1995) ("We believe that US West may be put on notice if it learns that the perpetrator has practiced widespread sexual harassment in the office place, even though US West may not have known that this particular plaintiff was one of the perpetrator's victims.") *and* Ferris v. Delta Air Lines, Inc., 277 F.3d 128, 136 (2d Cir. 2001) ("Delta had notice of Young's proclivity to rape co-workers. The fact that Young's prior rapes were not of Ferris but of other co-workers is not preclusive."), *with* Longstreet v. Ill. Dep't of Corr., 276 F.3d 379, 382 (7th Cir. 2002) (refusing to adopt a duty-to-warn standard and reasoning it would be unjust "that a response which seemed to be within the realm of reasonableness in one situation can, if ultimately it did not have the proper deterrent effect, be the sole basis for liability in another case even if the employer's response in the second case was clearly sufficient.").

Pauline, the employer would be obligated to not only investigate, but also correct the aberrant behavior.

B. Duty to Investigate

But assume no court ever expands an employer's preventative duties. Instead, something a bit more plausible occurs; employers adopt a company-wide monitoring system for more selfish reasons. Lawsuits are expensive, and the company prefers to nip any potential issues in the bud before a complaint is filed.[147] Or perhaps the company expands surveillance to protect its proprietary information.[148] Could that trigger any affirmative obligations for the employer?

There is no clear answer, but—depending on the information the AI reviews and the reports it sends to management—an employer might be obligated to investigate. Employers are duty-bound "to take prompt corrective action that is 'reasonably calculated to end the harassment'" once they receive notice of such behavior.[149] Currently, both actual and constructive notice demands harassment occur before any duty is triggered. Actual notice manifests once a report is filed, while constructive notice exists when harassment "takes place so openly and frequently that the employer would be expected to have observed the conduct."[150]

---

147. Courts are also willing to admit evidence of harassing comments about a plaintiff, even when the plaintiff is not present to hear them. *See, e.g.*, Schwapp v. Town of Avon, 118 F.3d 106, 111 (2d Cir. 1997) ("The district court also erred in failing to consider the eight additional [racially-charged] incidents that did not occur in Schwapp's presence."); Hurley v. Atlantic City Police Dep't, 174 F.3d 95, 110 (3d Cir. 1999) ("a plaintiff may show that, while she was not personally subjected to harassing conduct, her working conditions were nevertheless altered as a result of witnessing a defendant's hostility towards other women at the workplace."); Hawkins v. Anheuser-Busch, Inc., 517 F.3d 321, 336 (6th Cir. 2008) ("This court's caselaw therefore makes clear that the factfinder may consider similar acts of harassment of which a plaintiff becomes aware during the course of his or her employment, even if the harassing acts were directed at others or occurred outside of the plaintiff's presence."). This evidentiary problem also incentivizes employers to fully monitor all chatter.

148. A system, not unlike the one described for sexual harassment, is being used for this exact purpose. Adam K. Levin, *You're Being Watched at Work*, CHI. TRIB. (Apr. 16, 2018), http://www.chicagotribune.com/business/success/tca-you-re-being-watched-at-work-20180416-story.html [https://perma.cc/49QU-EL56].

149. Swenson v. Potter, 271 F.3d 1184, 1192 (9th Cir. 2001); *see also* Burlington Indus. v. Ellerth, 524 U.S. 742, 765 (1998); Faragher v. City of Boca Raton, 524 U.S. 775, 807 (1998).

150. J. Hoult Verkerke, *Notice Liability in Employment Discrimination Law*, 81 VA. L. REV. 273, 317 (1995); *see also* Hrobowski v. Worthington Steel Co., 358 F.3d 473, 478 (7th Cir. 2004) ("However, an employer could be charged with constructive notice where the harassment was sufficiently obvious.").

Courts seem willing to stretch the 'open and frequent' element of constructive notice when an employer monitors its employee's digital activity. The New Jersey Superior Court held in *Blakey v. Continental Airlines* that employers are not obligated to monitor e-mails, but "may not disregard the posting of offensive messages on company or state agency e-mail systems when the employer is made aware of those messages."[151]

This obligation has expanded slightly post-*Blakey*. In *Doe v. XYC Corp.*, the network administrator at a company that regularly monitored network activity for inappropriate behavior discovered an employee was visiting pornographic sites.[152] The department performed a limited investigation but never reported the behavior.[153] As it turned out, the employee was also visiting sites with child pornography, which the IT administrators could easily discover with a more extensive review.[154] But it was not until the employee abused a co-worker's child that anyone realized the extent of his behavior.[155] The defendant claimed they had no idea he was accessing child pornography.[156] However, the court imputed knowledge to the employer because "an investigation" of his illicit browsing "would have readily uncovered the full scope of Employee's activities."[157] Since the company had constructive knowledge of this activity, it was obligated to report the behavior to the police and either terminate the employee or take some other "internal action to stop [the employee's] activities."[158]

The old axiom that "hard cases make bad law" undoubtedly applies in this context. Viewing and creating child pornography is such a vile act that courts are more willing to push the boundaries of notice and duty. But there is a noticeable trend towards expanding the concept of notice when employers monitor an employee's activity. If someone who is duty-bound to report harassment to their employer sees—or could easily access a report

---

151. Blakey v. Cont'l Airlines, 751 A.2d 538, 551 (N.J. 2000) ("That does not mean that employers have a duty to monitor employees' mail. Grave privacy concerns are implicated . . . . It may mean that employers may not disregard the posting of offensive messages on company or state agency e-mail systems when the employer is made aware of those messages.").

152. Doe v. XYC Corp., 887 A.2d 1156, 1159 (N.J. Super. Ct. App. Div. 2005).

153. *Id.*

154. *Id.* at 1161.

155. *Id.* at 1161–62.

156. *Id.* at 1162.

157. *Id.* at 1167.

158. *Id.* at 1167–68.

identifying—potential problems, the court might find the employer on notice.[159] Notice rules "control incentives to invest in information."[160] Accurate information is worth investing in when it impacts the bottom-line.[161]

Under the current regime, investing in harassment-monitoring software is ill-advised since it might needlessly trigger investigations that would otherwise be unnecessary. However, as the cases out of New Jersey demonstrate, behavior both on and off a work terminal contribute equally to constructive notice. That creates an incentive to adopt a system that flags issues and permits corrective action. The risk-averse will want as comprehensive of a system as possible.

C. *The Thirst for Data*

Effective AI requires access to extraordinary amounts of data.[162] Without that access, it is nearly impossible to identify patterns and predict behavior. And employers are uniquely situated as they can institute massive systems of surveillance otherwise frowned upon when done by most Americans.[163] Courts do consider

---

159. Debord v. Mercy Health Sys. of Kan., Inc., 737 F.3d 642, 657 (10th Cir. 2013) ("Her Facebook post was not in accordance with Mercy's otherwise flexible reporting system for sexual harassment complaints, and the post, by itself, did not provide any notice to Mercy. Only when Weaver himself brought the post to [the HR Director's] attention did Mercy learn that, among many other complaints, Debord disliked Weaver's 'creepy hands.'"); *see also* Fisher v. Mermaid Manor Home for Adults, LLC, 192 F. Supp. 3d 323, 329–30 (E.D.N.Y. 2016) (noting an employer's failure to take remedial action after seeing an Instagram post that "compar[ed] Plaintiff to a fictional chimpanzee from the *Planet of the Apes* movie, created a hostile work environment" as the employer "knew of the harassment but did nothing about it."); Nischan v. Stratosphere Quality, LLC, 865 F.3d 922, 931 (7th Cir. 2017) ("Generally, for constructive notice to attach, the notice must 'come to the attention of someone who . . . has under the terms of his employment . . . a duty to pass on the information to someone within the company who has the power to do something about it.' . . . Once that person learns of the sexual harassment, the employer is considered to be on notice even if the victim never reported the harassment.").

160. Verkerke, *supra* note 150, at 319.

161. This idea was recently explored by Kate Klonick. Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018). Klonick concluded that companies tend to self-regulate and push for accurate content because obscenity, inaccuracies, and violence threaten "potential profits based in advertising revenue." *Id.* at 1627.

162. *See generally* GOODFELLOW ET AL., *supra* note 88 at 5–8 (noting that AI often need to be coded and experience their own learning to be most effective).

163. The Electronic Communications Privacy Act prohibits the interception of electronic data and communication. 18 U.S.C. § 2511 (2018). However, there is a carve-out for communications made on equipment furnished by a company to a user "in the ordinary course of its business." 18 U.S.C. § 2510(5)(a) (2018).

whether there is a reasonable expectation of privacy, and the method and manner of the invasion.[164] But generally, as long as the employee is on notice and the business either owns the data or has an agreement permitting access to data held off-site, the only problems that stand in the way of permitting access are potential rather than concrete.

i. Privacy Problems

As mentioned in Section III, a predictive AI could potentially access publicly available information posted by employees. But there are a host of privacy concerns in such a program and, as there is an implicit right to privacy in the U.S. Constitution, courts may be wary about such an intrusive act.[165] Currently, nothing prohibits employers from trawling public information. A handful of states statutorily restrict employers from requesting social media passwords,[166] but generally, any information posted online is deemed public.

Courts are typically unwilling to protect information in which there is no reasonable expectation of privacy.[167] And once you release information publicly, you cannot reasonably expect it remain private.[168] But, as the Supreme Court recognized in *United States v. Jones*, "[i]n the pre-computer age, the greatest protections of privacy were neither constitutional nor statutory, but practical."[169] Now, "many ordinary Americans are choosing to make

---

164. City of Ontario v. Quon, 560 U.S. 746, 762 (2010) ("[T]he extent of an expectation [of privacy] is relevant to assessing whether the search was too intrusive.").

165. *Cf.* Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890) (discussing how the courts respond to claims of breach of privacy).

166. *See, e.g.*, CAL. LAB. CODE § 980 (2019); CONN. GEN. STAT. ANN. § 31-40x (2019); 820 ILL. COMP. STAT. ANN. § 55/10 (2019); NEV. REV. STAT. ANN. § 613.135 (2019); VA. CODE ANN. § 40.1-28.7:5 (2019).

167. *See* United States v. Morel, 922 F.3d 1, 10 (1st Cir. 2019) *and* Katz v. United States, 389 U.S. 347, 351 (1967).

168. *Katz*, 389 U.S. at 351 ("What a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection."). The First Circuit Court of Appeals recently affirmed a conviction for possession of child pornography based on photos the defendant posted to an online image board. United States v. Morel, 922 F.3d 1 (1st Cir. 2019). Morel argued that he had a reasonable expectation of privacy as to the images he uploaded, and the Government thus violated the Fourth Amendment by reviewing the images without a warrant. *Id.* at 7–8. The Court disagreed, finding Morel had no "reasonable expectation of privacy in the images uploaded to Imgur" as "'everyone in the world can see' images uploaded to public Imgur albums, and that those images are available on Imgur's public galleries." *Id.* at 10.

169. United States v. Jones, 565 U.S. 400, 429 (2012) (Alito, J., concurring).

public much information that was seldom revealed to outsiders just a few decades ago," but fail to realize the implications of that behavior.[170] Something as innocuous as where someone ate lunch, or what they wore to the beach, "can lead to insights of a deeply sensitive nature," thanks to pattern recognition and AI.[171]

Of course, these Fourth Amendment cases do not constrain private employers.[172] Rather, civil liability in tort law typically reins in abusive behavior.[173] But a similar type of analysis applies in this context, as courts normally examine whether "a plaintiff reasonably expected that information about himself would remain 'private' after" disclosing it to other people.[174] Though highly underdeveloped, the cases examining this issue typically turn on the affirmative acts the plaintiff-employee took to protect the posted information.[175]

In *Ehling v. Monmouth-Ocean*, for example, a nurse sued her ex-employer for invasion of privacy after the hospital terminated

---

170. Riley v. California, 573 U.S. 373, 408 (2014) (Alito, J., concurring).

171. Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 420–21 (2017).

172. Some courts have analogized a party's reasonable expectation of privacy enforceable against a private employer to the right to be free from unreasonable searches and seizures under the Fourth Amendment. *See, e.g.*, Twigg v. Hercules Corp., 185 W. Va. 155, 159 (1990). But, generally, there is reticence to "transfer the jurisprudence of the cases involving government employers to actions against private employers." Borse v. Piece Goods Shop, Inc., 963 F.2d 611, 625 (3d Cir. 1992); *see also* Stein v. Davidson Hotel Co., 945 S.W.2d 714, 718 (Tenn. 1997) (collecting case law holding "that constitutional guarantees restrain government conduct and generally do not restrain the conduct of private individuals."); Whye v. Concentra Health Servs., 2013 U.S. Dist. LEXIS 137142, at *59 (D. Md. Sep. 24, 2013) (rejecting an invasion of privacy argument based on Fourth Amendment jurisprudence after noting that "[t]he Fourth Amendment does not apply to a search or seizure, even an arbitrary one, effected by a private party on his own initiative.") (alteration in original) (quoting Skinner v. Ry. Labor Execs. Ass'n, 489 U.S. 602, 614 (1989)).

173. Additionally, some states also have statutory protections for personal and private information. For example, California's Consumer Privacy Act (CCPA), effective January 1, 2020, may require employers who adopt a system that combs and collects data about its employees to put those employees on notice of the practice. CAL. CIV. CODE § 1798.100 (2019).

174. Lior Jacob Strahilevitz, *A Social Networks Theory of Privacy*, 72 U. CHI. L. REV. 919, 921 (2005).

175. *See, e.g.*, Meyers v. Siddons-Martin Emergency Grp. LLC, No. 16-1197, 2016 WL 5337959, at *3–*4 (E.D. La. Sep. 23, 2016) (dismissing an invasion of privacy claim brought against the plaintiff's former employer after the plaintiff was terminated for offensive Facebook posts discovered after the employee logged into his account "on a computer that was owned by Siddons-Martin," as that act removed any "reasonable expectation of privacy" he had over his posts); Ehling v. Monmouth-Ocean Hosp. Serv. Corp., 872 F. Supp. 2d 369, 374 (D.N.J. 2012).

her, in part, for a controversial statement she made on Facebook.[176] Ehling premised her claim on the fact the post was not public; rather, anyone who was not "invited to be her Facebook 'friend,' . . . could not access and view postings on [her] Facebook 'wall.'"[177] Monmouth-Ocean moved to dismiss, arguing anyone she added as a Friend could read the post, and thus she never expected it to remain private.[178] The court denied the motion, finding the fact Ehling "actively took steps to protect her Facebook page from public viewing" shows she could have "had a reasonable expectation that her Facebook posting would remain private."[179] However, courts have also found even the simple act of posting waives any expectation of privacy, as there is always a possibility those within the protected space could share the information with others.[180] As the *Ehling* court notes, "reasonableness (and offensiveness) are highly fact-sensitive inquiries."[181]

What analysts can glean from the case law, however, is that an employee who expects information posted publicly—i.e. not just accessible to friends and followers, but anyone with an internet connection—to remain private is living in a fantasy. But trying to peek behind the curtain separating the public from a user's limited world of friends and followers might be a step too far.

ii. The Cost of Litigation and Fear

The system this article envisions is supposed to warn; not accuse. But that distinction gets murky when straddling the line

---

176. *Ehling*, 872 F. Supp. 2d at 369–370.
177. *Id.* at 370.
178. *Id.* at 372.
179. *Id.* at 374.
180. *See, e.g.*, Smyth v. Pillsbury Co., 914 F. Supp. 97, 101 (E.D. Pa. 1996) ("[W]e do not find a reasonable expectation of privacy in e-mail communications voluntarily made by an employee to his supervisor over the company e-mail system notwithstanding any assurances that such communications would not be intercepted by management. Once plaintiff communicated the alleged unprofessional comments to a second person (his supervisor) over an e-mail system which was apparently utilized by the entire company, any reasonable expectation of privacy was lost."); Heldt v. Guardian Life Ins. Co. of Am., No. 16-cv-885-BAS-NLS, 2019 U.S. Dist. LEXIS 25315, at *18 (S.D. Cal. Feb. 15, 2019) (granting summary judgment on an invasion of privacy claim brought against a private insurer who disclosed information gleaned from the plaintiff's social media page to an insurance investigator as the plaintiff "voluntarily shared the information with his Facebook friends knowing there is a possibility that his friends could share the information with others," and thus "did not have a reasonable expectation of privacy in that information.").
181. *Ehling*, 872 F. Supp. 2d at 374.

between truth and probability. Interestingly, leaked data alone should not cause any problems. The data fed into this system is, again, public facing.[182] Anyone with an internet connection could find and review that information; the system just aggregates it into one place.[183] Rather, the 'red-flag' or warning itself is what threatens the most harm to someone's reputation if leaked.

Defamation claims are the common road to redress reputational harm. The warning—"Harvey has a 90% likelihood of sexually harassing Sally"—seems to state, or at least imply, a fact.[184] Truth remains the absolute defense to defamation, though it is difficult to argue a warning or probability is entirely truthful. And, while you could explain how the probability was reached based on true information, doing so requires cracking the proprietary black box protecting algorithmic programs. Very few companies would sacrifice their secrets to help defend against a defamation suit brought against one client.[185]

It is more likely that—rather than argue truth—employers would instead rely on the qualified "common interest" privilege[186] that normally helps sink the (admittedly rare) defamation suits

---

182. Problems could arise, of course, if the information used by the system is not public. There would be some interplay between an employee's private information, as maintained by human resources, and the public facing information compiled for the system's analysis. And some states mandate that businesses protect their employees' private information. In California, for example, employers must "implement and maintain reasonable security procedures and practices" to safeguard its employees' "personal information." CAL. CIV. CODE § 1798.81.5 (2019).

183. As discussed *supra* Section IV.C.i, employees face an uphill battle if arguing information they posted publicly on social media must remain private.

184. Even if it is more opinion than fact, "[a] defamatory communication may consist of a statement in the form of an opinion . . . if it implies the allegation of undisclosed defamatory facts as the basis for the opinion." RESTATEMENT (SECOND) OF TORTS § 566 (1979). Unless the public knows and understands how an early-warning system like the one envisioned by this article works, they could reasonably assume the probability is based on prior acts of harassments, which may or may not exist.

185. This issue has arisen already with programs used for criminal sentencing. Eric Loomis was sentenced to six years in prison because "[COMPAS]," a proprietary algorithm used in sentencing, believed he had a high risk of recidivism. *See* Adam Liptak, *Sent to Prison by a Software Program's Secret Algorithms*, N.Y. TIMES (May 1, 2017), https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-softwa re-programs-secret-algorithms.html [https://perma.cc/Q29M-FQS2]. The Wisconsin Supreme Court upheld the sentence, reasoning that while Loomis "cannot review and challenge how the COMPAS algorithm calculates risk, he can at least review and challenge the resulting risk scores set forth in the report." State v. Loomis, 881 N.W.2d 749, 761 (Wis. 2016).

186. Martha S. West, *Preventing Sexual Harassment: The Federal Court's Wake-Up Call for Women*, 68 BROOKLYN L. REV. 457, 509 (2002).

brought by accused-harassers. While defamation law varies from state-to-state, that privilege generally protects intra-company communications between employees who share a common interest in the subject, such as a human resources department concerned about sexual harassment.[187] A warning spat out by the system to human resources likely falls in the realm of intra-company messaging.[188] However, employers lose that privilege when they 'excessively' publish the communication.[189] And that is where a potential leak could be hazardous. Recognizing that some leaks are inevitable, courts examining an excessive publication challenge tend to review what an employer did to keep the alleged defamatory information within the proper channels.[190]

Inevitably, creative plaintiffs will find other avenues for vindicating reputational harm caused by an early-warning system. The Federal Fair Credit Reporting Act (FCRA), despite its title, covers more than credit, encompassing "any written, oral, or other communication of any information by a consumer reporting agency bearing on a consumer's . . . character, general reputation, [or]

187. *Id.* at 508 (exploring the effect of a qualified privilege, as well as other defenses, on defamation suits rooted in sexual harassment accusations within various jurisdictions).

188. California has even codified the qualified privilege if the communication pertains to alleged sexual harassment. CAL. CIV. CODE § 47(c) (2019) (explaining the privilege applies "to and includes a complaint of sexual harassment," as well as a prior employer's "decision to not rehire" an employee "based upon the employer's determination that the former employee engaged in sexual harassment.").

189. West, *supra* note 186, at 510 ("Another possible way for a terminated employee to prove abuse of the privilege protecting an employer's workplace communications is to prove 'excessive publication'—that the employer distributed the allegedly defamatory information outside the workplace.").

190. Garziano v. E. I. Du Pont de Nemours & Co., 818 F.2d 380, 393 (5th Cir. 1987) (remanding a case for the fact-finder to determine whether rumors about an employee's termination for alleged sexual harassment were "the product of supervisors discussing the contents of" an internal communication concerning the reasons for the termination "with persons outside the 'circle' of interested employees, or whether the rumors were only routine scuttlebutt due to Garziano's discharge."); Esmark Apparel v. James, 10 F.3d 1156, 1163 (5th Cir. 1994) (finding "minor leaks" caused by supervisors sharing the reasons the company terminated an employee with "their immediate families" did not evidence "that the company . . . abuse[d] its qualified privilege through excessive publication"); Tacka v. Georgetown Univ., 193 F. Supp. 2d 43, 51 (D.C. 2001) ("[Q]ualified privilege may be lost . . . 'if the publication occurs outside normal channels, is otherwise excessive, or was made with malicious intent.'") (quoting District of Columbia v. Thompson, 570 A.2d 277, 292 (D.C. 1990)); White v. Blue Cross & Blue Shield of Mass. 809 N.E.2d 1034, 1038 (2004) (holding that the qualified privilege protecting intra-company communication is "lost only when the employer recklessly makes 'unnecessary, unreasonable or excessive' publications").

personal characteristics" collected for "employment purposes."[191] That includes "evaluating a consumer for employment, promotion, reassignment or retention as an employee."[192]

Indeed, some have already challenged broad 'risk assessments' collected for, and produced to, their employers. A technician who worked for Superior, a third-party installation company, brought one such claim in *Ernst v. Dish Network, LLC*.[193] Dish Network, the named-defendant, contracted to have Superior technicians install satellite dishes for Dish's customers.[194] As part of this arrangement, Dish required that Superior run criminal background checks on their employees and forward Dish a "summary report" with "the individual's risk rating" denoted as "'high risk,' 'low risk,' or 'review.'"[195] "High risk" employees, like Ernst, were forbidden from installing equipment.[196] He sued Dish, arguing he had a right to inspect and correct the report under the FCRA.[197] The court concluded "[a] 'high risk' rating on the Summary Report in effect says that [Plaintiff] . . . has done something highly improper that impugns his moral character."[198] And because Ernst's employer "used the information in the Summary Report . . . for 'reassignment or retention as an employee,'" it was subject to the FCRA.[199]

As *Ernst* shows, there is no defense to an FCRA action just because another entity compiled the report, or the report just summarizes the information actually collected. Any adopters of an early-warning system should thus strongly consider producing a copy of any reports generated to impacted employees. Unfortunately, this practice could result in higher amounts of self-censorship as people try to 'correct' their ranking by meeting with fewer opposite sex colleagues, or avoiding mentorship-relationships

---

191. 15 U.S.C. § 1681a(d)(1) (2018). Some states have adopted their own versions of the FCRA. California's Investigative Consumer Reporting Agencies Act (ICRAA) requires anyone evaluating publicly-available information concerning a consumer's "character, general reputation, personnel characteristics, or mode of living, for employment purposes" to disclose such information "within seven days after receipt," even when the entity forgoes "the services of an investigative consumer reporting agency." CAL. CIV. CODE § 1786.53 (2019).

192. 15 U.S.C. § 1681a(h) (2018).

193. Ernst v. Dish Network, LLC, 49 F. Supp. 3d 377, 378 (S.D.N.Y. 2014).

194. *Id.* at 379.

195. *Id.*

196. *Id.*

197. *Id.* at 378.

198. *Id.* at 382.

199. *Id.* at 384 (quoting 15 U.S.C. § 1681a(h) (2018)).

with their subordinates.[200] The risk of litigation and ensuing side effects of a massive surveillance effort may prove too high for implementation. Only time will tell.

V. "Looks Like You've Had a Bit Too Much to Think!": Conclusions

"We shall meet in the place where there is no darkness."[201]

Consider the following an example. There will be, in the not so distant future, a man named Harvey.[202] Harvey will run a fairly successful public relations firm with lots of employees, including a young lady named Sally. Sally will be twenty-one years old, fresh out of college, and dream of running her own PR firm. Harvey will take a liking to Sally and offer her an internship. He lets his Human Resources director Jim know about the new hire and provides a basic rundown of her job—Sally will shadow Harvey, watch him run everything, and generally help out around the office. Jim puts these variables into 'The Eye;' a handy system that generally monitors employee activity and flags potential problems for Jim to review.

At first, there are no real issues—'The Eye' flags the power differential between Harvey and Sally, but only predicts a 15% chance of harassment. However, it began noticing some odd behavior from Harvey's work-terminal. Her name kept appearing in his search history; they tend to clock in and out at around the same time, even when Harvey stays well past closing; and recently he started interacting with images Sally posted years ago wherein she is wearing revealing clothes.[203] It sends Jim a warning—Harvey is now a high-risk harasser. Jim prints out the report and calls Harvey into his office. He explains how sensitive these matters are, and lets Harvey know 'The Eye' thinks there may be a problem. He provides Harvey a copy of his report and cautions him to stay away

---

200. Steve Hendrix, Ellie Silverman & Marc Fisher, *#MeToo has a 'Chilling Effect' on Workplace Camaraderie*, CHI. TRIB. (Jan. 28, 2018), http://www.chicagotribune .com/business/ct-metoo-workplace-camaraderie-20180128-story.html [https://perma .cc/Z5EP-XQ5U]; *see also* Vicki Schultz, *The Sanitized Workplace*, 112 YALE L.J. 2061 (2003) (discussing the effect of sex in the workplaces).

201. GEORGE ORWELL, NINETEEN EIGHTY-FOUR 25 (Harcourt Inc. 1977) (1949).

202. Names, characters, businesses, places, events, locales, and incidents are purely products of the author's imagination. Any resemblance to eventual persons, living or dead, or eventual events is entirely coincidental.

203. This feature sold Jim; AI analyzes pixel by pixel, compares it to publicly available images on the internet of similar skin hue, determines to 99% probability how many pixels are of skin and how many are clothes, then assesses risk based on how many pictures are being liked of someone half-naked.

from Sally. Harvey is annoyed by this—Sally was a nice girl and he thought they had a real connection—but he agrees and heads back to his office.

Sally notices Harvey seemed a bit more stand-offish than usual, so she decides to help out by tidying up early the next morning. While sorting the paperwork in his office, Sally sees the report—'Harvey, CEO; 93% Risk Against Sally; Intervention MANDATORY.' Sally could not believe her eyes. She felt betrayed; Harvey wanted to harass her! She took a picture of the report and sent it to her friend Jenny, a journalist working on a report exploring sexual harassment in the modern workplace. Jenny posts a copy of the report to her blog, and tags everyone she can to out Harvey as a harasser.

Sexual harassment law is failing the harassed. It was designed to prevent harassment but cannot protect a victim until they are actually harmed. As this paper explored, artificial intelligence could be the key to this problem. As long as the program is tailored to consider subjective experience, a company could step in and prevent harm before it ever occurs. Such a system could even capture, organize, and centralize all of the rumors and complaints that swirl around a company.

The tricky parts are what employers do with the information and how employees view the system. Employers should not take any adverse action unless harassment actually occurs. These warnings should instead simply prompt an investigation and discussion. Instead of seeing these reports or scores as accusations, they should be framed the same as credit scores. Having a low credit score does not make you a bad person—it simply changes your risk profile. And once you know your score is low, either because you were called into human resources or you received a copy of the report, you can work on making it better. Small, incremental adjustments in attitude and perception can go a long way in promoting empathy and cutting back on harassment.