

# Engagement in Deepfake Pornography

---

**AUTHOR**

Nicola Döring, Thuy Dung Le

---

**KEYWORDS**

*AI-generated pornography, deepfakes, AI-generated Image-Based Sexual Abuse (AI-IBSA), AI-generated Non-consensual Intimate Images (AI-NCII)*

---

**ABSTRACT**

In the age of AI, the production of pornography has changed in two key ways: (a) purely synthetic images can be generated without any resemblance to real persons, and (b) deepfake pornography of any individual can be created (Döring et al., 2025b). Deepfake pornography refers to sexually explicit images or videos in which a person's face or likeness is superimposed onto pre-existing pornographic content, using AI-based face-swapping or image synthesis technology (Döring et al., 2025a). In most cases, this occurs without the consent of the depicted person and without the consent of the pornography performers whose material is appropriated. While the term deepfake pornography is widely used in public and academic discourse, more precise terms that foreground the nonconsensual nature of this practice are AI-generated nonconsensual intimate images (AI-NCII) and AI-generated image-based sexual abuse (AI-IBSA). A growing number of countries have enacted (e.g., Italy, the UK) or plan to enact (e.g., Germany) laws criminalizing the production and distribution of such nonconsensual intimate media content that depicts both private individuals — including minors — and public figures such as politicians, journalists, and performing artists. The harm of nonconsensual deepfake pornography is threefold: (1) it violates the right to privacy and personal integrity by undermining data protection, reputation, and psychological autonomy; (2) its strategic use to shame and defame women journalists, activists, and politicians threatens free-

dom of expression and gender equality in public and political participation; (3) it impairs the economic and moral rights of pornography performers whose work and bodies are exploited without authorization or compensation (Yavuz, 2025). On social media, considerable debate surrounds this phenomenon, ranging from potential perpetrators exchanging tips on how to generate such material, to bystanders seeking information on where to find or how to combat it, and victims searching for support (Döring et al., 2025a; Le & Döring, 2026). This entry provides a set of variables for identifying different engagement roles — perpetrator, victim, and bystander — as they appear in online discussion posts.

---

**FIELD OF APPLICATION/THEORETICAL FOUNDATION**

Deepfake pornography, also referred to as AI-generated image-based sexual abuse (AI-IBSA) or AI-generated non-consensual intimate images (AI-NCII), is a form of image-based digital violence that extends sexual and gender-based violence into the digital sphere (Döring et al., 2025a). Accordingly, theoretical perspectives on engagement in nonconsensual deepfake pornography require a gender-based approach. Gender role theory (Eagly, 1987), sexual script theory (Simon & Gagnon, 1984), social dominance theory (Pratto & Stewart, 2011), and objectification theory (Fredrickson & Roberts, 1997) help explain why men are more likely to engage in the creation and consumption of AI-IBSA, while targeted women are more likely to experience social, professional, economic, and psychological harm. Women who are targeted with deepfake pornography face a disproportionate burden compared to men because their sexuality is socially policed through sexual double standards and slut-shaming norms, meaning that a fabricated sexual image activates pre-existing mechanisms of reputational destruction, ostracism, and silencing that have no equivalent for



<https://doi.org/10.34778/mm485x54>

© 2026, the authors. This work is licensed under the “Creative Commons Attribution – NonCommercial – NoDerivatives 4.0 International” license (CC BY-NC-ND 4.0)



male targets. Several empirical studies have explored engagement with deepfake pornography. Based on large-scale online surveys, Umbach et al. (2026) found that 3.2% of adults across three countries had created, distributed, or threatened to distribute AI-IBSA, while 18.0% reported viewing such content. Umbach et al. (2024) identified a 2.2% victimization rate of AI-NCII across ten countries. Across studies, men and younger individuals consistently show higher rates of perpetration, whereas women disproportionately report emotional, psychological, and professional harms.

To better understand these dynamics, recent studies apply the participant role approach (Salmivalli, 1999) and routine activity theory (Cohen & Felson, 1979), distinguishing between perpetrators, victims, and bystanders. Perpetrators create or distribute deepfake pornography, victims are depicted without consent, and bystanders encounter such material without directly participating in the abuse (Le & Döring, 2026). Mirroring above-cited survey results, social media analyses show that perpetrators posting about the issue are primarily identified as men, while victims and bystanders are predominantly women, often within close social networks of perpetrators. Experiences associated with these roles range from perpetrators publicly discussing the technicalities and the joys of creating such material to bystanders seeking intervention information and victims searching for support (Döring et al., 2025a; Le & Döring, 2026). Studies underscore the multiple harms of nonconsensual deepfake pornography and the need for role-specific interventions, including stronger platform governance, digital ethics education, bystander intervention programs, trauma-informed support services, and legal protections addressing AI-IBSA.

---

#### **REFERENCES/COMBINATION WITH OTHER METHODS OF DATA COLLECTION**

Social media content analyses (e.g., of Reddit, X/Twitter, and Facebook), as well as analyses of other materials such as court documents (Witt et al., 2024), provide relevant insights into people's experiences with deepfake pornography across different participant roles. Such content analysis studies can be conducted as qualitative analyses (e.g., Popova, 2020), manual quantitative analyses (e.g., Döring et al., 2025a), or com-

putational analyses (e.g., Han et al., 2024). Some studies combine qualitative and quantitative content analysis methods to capture both large-scale patterns and nuanced experiences (e.g., Le & Döring, 2026; Timmerman et al., 2023). Content analysis approaches can also be combined with oral interviews (e.g., Flynn et al., 2025) and written surveys (e.g., Umbach et al., 2026) involving victims, perpetrators, and bystanders of deepfake pornography to explore their experiences, motivations, and behavioral consequences. Some studies have additionally used experimental designs that present descriptions of different deepfake pornography cases as stimulus material (e.g., varying the gender, age, or celebrity status of victims) to examine how specific case characteristics influence empathy, ethical concern, willingness to help, or victim blaming among observers (e.g., Fido et al., 2022; Kugler & Pace, 2021; Henry et al., 2026).

---

#### **EXAMPLE STUDIES FOR MANUAL QUANTITATIVE CONTENT ANALYSES**

Several content analysis studies have investigated people's experiences with deepfake pornography across different participant roles. These include a computational content analysis of the sexual deepfake marketplace MrDeepFakes, analyzing more than 42,000 deepfake videos, 8,000 forum threads, 43,000 forum posts, and 611,000 public user profiles (Han et al., 2024); a manual quantitative content analysis of 390 English-language public Reddit posts related to AI pornography experiences (Döring et al., 2025a); and a qualitative content analysis of videos, metadata, and paratexts from two online deepfake communities, Mrdeepfakes.com and v/DeepFake (Popova, 2020).

The current example study (Le & Döring, 2026) uses a combination of manual quantitative and qualitative content analysis, anchored in Routine Activity Theory (RAT; Cohen & Felson, 1979). The manual quantitative content analysis is based on a sample of 134 Reddit initial posts published between 2017 and 2025. The codebook was developed deductively from the research questions, prior data, and existing studies (Powell et al., 2019; Wei et al., 2024; Statistics Canada, 2024), and was further refined inductively using example material. The codebook was evaluated for its three key quality criteria: objectivity, validity, and reliability (Döring, 2023): To ensure objectiv-



ity, all concepts are explained in detail with coding instructions and examples, and the full codebook is publicly available (<https://osf.io/hcnja/>). To ensure validity, all categories in the codebook are grounded in the theoretical and empirical literature with respective references. To ensure reliability, all categories were tested for intercoder reliability between two independent trained coders using a sub-sample of 40 Reddit initial posts. A codebook excerpt with concept definitions, operationalizations, reliabil-

ity coefficients, and sources can be found below (see Table 1). To analyze online discussions about deepfake pornography, one can (a) filter the material to only include threads and posts that address specific cases and clearly reveal the involvement roles of posters, or (b) skip prefiltering and include posts with unclear connection to specific cases and involvement roles which then requires the addition of an “unclear involvement role” category.

**Table 1.** Codebook to Identify Engagement Roles in Nonconsensual Deepfake Pornography Based on Online Posts

<b>Engagement Roles in Nonconsensual Deepfake Pornography as Expressed in Online Posts</b>	<b>Operationalizations</b>	<b>Reliability (Cohen’s Kappa coefficient)</b>	<b>Source</b>
(1) Perpetrator: a person who creates, distributes, and/or threatens to share deepfake pornography of another person without the depicted person’s and the pornography performers’ consent.	Perpetrator: code 1 = yes / code 0 = no Example quote from online post: “I tried to generate nude images with an AI app, but I had several problems.”	1.00	Flynn et al., (2022); Umbach et al. (2026)
(2) Victim: a private person or public figure whose likeness is depicted in, whose images are shared in, or who is threatened with the circulation of deepfake pornography without their consent.	Victim: code 1 = yes / code 0 = no Example quote from online post: “A child posted a deepfake porn video with my face on 4chan.”	.89	Flynn et al., (2022); Umbach et al. (2026)
(3) Bystander: a person who becomes aware of nonconsensual deepfake pornography practices—whether by viewing, witnessing, discovering, or being shown the material—without directly participating in them.	Bystander: code 1 = yes / code 0 = no Example quote from online post: “I discovered two months ago that my boyfriend was creating deepfake pornography of my friends and family.”	.90	Mainwaring et al. (2024); Scott et al. (2022)
(4) Unclear involvement role: a person whose involvement role in specific cases of deepfake pornography practice is not made explicit in the post.	Unclear involvement role: code 1 = yes / code 0 = no Example quote from online post: “I can’t believe how easy it has become to create this kind of content with AI.”	-	-

For a manual quantitative content analysis of online discussions about deepfake pornography, the engagement role variables introduced above can be combined with variables that characterize specific cases discussed (e.g., the relationship between victim and perpetrator or between bystander and victim; gender and age of the parties involved; Le & Döring, 2026). Furthermore, subtypes of perpetrators, victims, and bystanders can be differentiated based on various behaviors related to deepfake pornography. For example, some perpetrators target private persons from their social circle, others public figures, some act alone, others act within a community of perpetrators. Some bystanders try to help victims,

others blame victims and defend perpetrators. Some victims experience and fight single incidents of victimization, others multiple and prolonged ones as explained in their online posts. It is important to note that in each case of nonconsensual deepfake pornography there is always a second type of victim, namely the pornography performers whose material is appropriated without their consent. However, their voices and perspectives are so far largely absent in academic and public debates (Döring et al., 2025b). To include them, one would need to find online discussion forums where pornography performers are active and code the performers as a specific victim subtype.

## REFERENCES

- Cohen, L. E., & Felson, M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review*, 44(4), 588. <https://doi.org/10.2307/2094589>
- Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften [Research Methods and Evaluation in the Social and Human Sciences]* (6th ed.). Springer.
- Döring, N., Le, T. D., & Miller, D. J. (2025a). Experiences with AI-Generated Pornography: A Quantitative Content Analysis of Reddit Posts. *Archives of Sexual Behavior*. <https://doi.org/10.1007/s10508-025-03227-x>
- Döring, N., Le, T. D., Vowels, L. M., Vowels, M. J., & Marcantonio, T. L. (2025b). The Impact of Artificial Intelligence on Human Sexuality: A Five-Year Literature Review 2020–2024. *Current Sexual Health Reports*, 17(1), 4. <https://doi.org/10.1007/s11930-024-00397-y>
- Eagly, A. H. (1987). Sex differences in social behavior: A social-role interpretation. Lawrence Erlbaum Associates, Inc.
- Fido, D., Rao, J., & Harper, C. A. (2022). Celebrity status, sex, and variation in psychopathy predict judgments of and proclivity to generate and distribute deepfake pornography. *Computers in Human Behavior*, 129, 107141. <https://doi.org/10.1016/j.chb.2021.107141>
- Flynn, A., Powell, A., Eaton, A., & Scott, A. J. (2025). Sexualized Deepfake Abuse: Perpetrator and Victim Perspectives on the Motivations and Forms of Non-Consensually Created and Shared Sexualized Deepfake Imagery. *Journal of Interpersonal Violence*, 08862605251368834. <https://doi.org/10.1177/08862605251368834>
- Flynn, A., Powell, A., Scott, A. J., & Cama, E. (2022). Deepfakes and Digitally Altered Image Abuse: A Cross-Country Exploration of an Emerging Form of Image-Based Sexual Abuse. *The British Journal of Criminology*, 62(6), 1341–1358. <https://doi.org/10.1093/bjc/azab111>
- Fredrickson, B. L., & Roberts, T.-A. (1997). Objectification Theory: Toward Understanding Women’s Lived Experiences and Mental Health Risks. *Psychology of Women Quarterly*, 21(2), 173–206. <https://doi.org/10.1111/j.1471-6402.1997.tb00108.x>
- Han, C., Li, A., Kumar, D., & Durumeric, Z. (2024). Characterizing the MrDeepFakes Sexual Deepfake Marketplace (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2410.11100>
- Henry, N., Umbach, R., Shelby, R., Beard, G., & Given, L. M. (2026). ‘It’s still abuse’: Community attitudes and perceptions on AI-generated image-based sexual abuse. *Information, Communication & Society*, 1–21. <https://doi.org/10.1080/1369118X.2026.2613437>
- Kugler, M. B., & Pace, C. (2021). Deepfake Privacy: Attitudes and Regulation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3781968>
- Le, T. D., & Döring, N. (2026). Perspectives on Non-consensual Deepfake Pornography: A Content Analysis of Reddit Discussions Initiated by Perpetrators, Victims, and Bystanders. *Sexuality & Culture*. <https://doi.org/10.1007/s12119-026-10566-x>

- Mainwaring, C., Scott, A. J., & Gabbert, F. (2024). Facilitators and Barriers of Bystander Intervention Intent in Image-Based Sexual Abuse Contexts: A Focus Group Study with a University Sample. *Journal of Interpersonal Violence*, 39(11–12), 2655–2686. <https://doi.org/10.1177/08862605231222452>
- Popova, M. (2020). Reading out of context: Pornographic deepfakes, celebrity and intimacy. *Porn Studies*, 7(4), 367–381. <https://doi.org/10.1080/23268743.2019.1675090>
- Powell, A., Henry, N., Flynn, A., & Scott, A. J. (2019). Image-based sexual abuse: The extent, nature, and predictors of perpetration in a community sample of Australian residents. *Computers in Human Behavior*, 92, 393–402. <https://doi.org/10.1016/j.chb.2018.11.009>
- Pratto, F., & Stewart, A. L. (2011). Social Dominance Theory. In D. J. Christie (Ed.), *The Encyclopedia of Peace Psychology* (1st ed.). Wiley. <https://doi.org/10.1002/9780470672532.wbep253>
- Salmivalli, C. (1999). Participant role approach to school bullying: Implications for interventions. *Journal of Adolescence*, 22(4), 453–459. <https://doi.org/10.1006/jado.1999.0239>
- Scott, A. J., Mainwaring, C., Flynn, A., Powell, A., & Henry, N. (2022). The extent and nature of image-based sexual abuse among Australian youths: Perspectives from victims, perpetrators and bystanders. In H. Lim (Ed.), *Interpersonal violence against children and youth* (pp. 85–108). Rowman & Littlefield Publishers.
- Simon, W., & Gagnon, J. H. (1984). Sexual scripts. *Society*, 22(1), 53–60. <https://doi.org/10.1007/BF02701260>
- Statistics Canada. (2024, October 1). Age Categories, Life Cycle Groupings. Statistics Canada. <https://www.statcan.gc.ca/en/concepts/definitions/age2>
- Timmerman, B., Mehta, P., Deb, P., Gallagher, K., Dolan-Gavitt, B., Garg, S., & Greenstadt, R. (2023). Studying the Online Deepfake Community. *Journal of Online Trust and Safety*, 2(1). <https://doi.org/10.54501/jots.v2i1.126>
- Umbach, R., Henry, N., Beard, G. F., & Berryessa, C. M. (2024). Non-Consensual Synthetic Intimate Imagery: Prevalence, Attitudes, and Knowledge in 10 Countries. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–20. <https://doi.org/10.1145/3613904.3642382>
- Umbach, R., Henry, N., Shelby, R., Stevens, G., & Gonzalez-Pons, K. (2026). AI-generated image-based sexual abuse: Perpetration and consumption across three regions. *Computers in Human Behavior*, 179, 108935. <https://doi.org/10.1016/j.chb.2026.108935>
- Wei, M., Consolvo, S., Kelley, P. G., Kohno, T., Matthews, T., Meiklejohn, S., Roesner, F., Shelby, R., Thomas, K., & Umbach, R. (2024). Understanding Help-Seeking and Help-Giving on Social Media for Image-Based Sexual Abuse. <https://doi.org/10.48550/ARXIV.2406.12161>
- Witt, A., Henry, N., & Beard, G. (2024). Judicial Constructions of Harm in Australian Image-Based Sexual Abuse Case Law: A Feminist Discourse Analysis. *Social & Legal Studies*, 33(6), 951–972. <https://doi.org/10.1177/09646639241233992>
- Yavuz, C. (2025). Adverse human rights impacts of dissemination of nonconsensual sexual deepfakes in the framework of European Convention on Human Rights: A victim-centered perspective. *Computer Law & Security Review*, 56, 106108. <https://doi.org/10.1016/j.clsr.2025.106108>