

Accepted Author Manuscript

To be published in 2025. Please refer to the website of [Transformative Works and Cultures](#) for the article's DOI

Special Issue “AI and Fandom”, Transformative Works and Culture

“This is a deepfake!”: Celebrity scandals, parodic deepfakes, and a critically speculative ethics of care for fandom research in the age of Artificial Intelligence

Eva Cheuk-Yin Li, Lancaster University¹

Ka-Wei Pang, The Chinese University of Hong Kong

[0.1] Abstract—Fans are increasingly aware of deepfake—believable AI-fabricated videos—and are therefore more skeptical of unverified information, even when visual evidence appears convincing. This article offers a methodological reflection on analyzing a deepfake event in which fans produced and circulated AI-generated disinformation to playfully undermine the credibility of a celebrity's video scandal. We explore the complex human-community-machine interactions (HCMI) between fans and AI-generated images, and we discuss how researchers can ethically (re)present their findings. We call for rethinking the “fan first” principle, a core tenet of ethical fandom research. Drawing on Puig de la Bellacasa's technoscientific theorization of care, we propose a critically speculative ethics of care in fandom research, guided by three principles: (1) thinking with fans, (2) thinking for fandom, and (3) thinking beyond fans and fandom. This approach is particularly relevant in a digital media ecology where generative AI and fan practices mutually transform each other. Our discussion also serves as a springboard for further explorations of ethics related to AI, including its impact on trust, social relations, and power in fandom research.

[0.2] Keywords—AI; Counter-information; Disinformation; Ethics of Care; Fans first; Fan studies ethics; Feminist ethics; Human-Community-Machine Interactions (HCMI); Machine learning; Parodic deepfake

¹ Email: eva.li@lancaster.ac.uk

1. Introduction

[1.1] The rise of accessible AI-powered image generation and synthesis technology in the 2020s has intensified the “Is it deepfake?” puzzle. Deepfakes—AI-generated face-swapping and voice synthesization (Karnouskos 2020, 138)—have made it increasingly difficult to distinguish genuine content from fabricated content. When beloved celebrities face photo or video scandals, fans often instinctively claim the content is “just a deepfake”. While deepfakes have undermined the credibility of visual media, they are also highly believable, making them effective tools for disinformation. A few months before Taylor Swift’s deepfake controversy in early 2024, a deepfake event occurred in the Thai Boys Love and Girls Love fandoms, where a celebrity video scandal opened Pandora’s box around the creation, dissemination, and (re)presentation of deepfakes in this age of generative AI (GenAI). Girls Love (GL) is a genre of queer media featuring emotional or physical intimacy between women, functioning similarly to Boys Love (BL) media, which focuses on male homoeroticism (Li and Pang 2024b). This article presents acafans’ methodological reflections on this event, which were not initially intended for publication. We begin by providing a brief overview of the context and our methodological approach. Then, we discuss fans’ responses to the Thai GL celebrity video scandal, particularly how some created parodic deepfakes to undermine the authenticity of the scandalous video. This led us to an ethical dilemma: how should researchers engage with fans’ creative practices that involve GenAI? Our reflections point towards a critically speculative ethics of care in fan studies amid a digital environment increasingly shaped by AI.

[1.2] Academic knowledge on AI has been “black-boxed” by predominantly focusing on computer science, technicalities, and codes, leaving space for social sciences and humanities to examine how machine learning reshapes our understanding of social relations and reality (Amoore et al. 2023, Browne et al. 2023). This article contributes to that exploration by offering a methodological reflection on a deepfake event in fandom. In response to a scandal involving an alleged deepfake video in the Thai BL and GL fandoms, fans created a defensive parodic deepfake to discredit the scandalous video. While this fannish practice was well-intentioned, it may have unintended and undesirable consequences. As researchers, we found ourselves caught between an ethical dilemma when seeking to (re)present the scandal and the parodic deepfake. Reflecting on these challenges, we argue that the “fans first” principle in fandom studies needs expansion in a digital ecology where GenAI and machine learning are transforming fannish practices. Building upon the framework of fandom as human-community-

machine interactions (HCMI) (Li and Pang 2024a), which emphasizes the interplay between human and nonhuman agents within collective social and cultural practices, we propose a critically speculative ethics of care. This approach repositions humanity within an increasingly networked digital media ecosystem shaped by GenAI. However, it is important to note that while research on fannish parodic practices inform academic knowledge about fans' creativity, it may also inadvertently overlap with toxic practices employed by anti-fans—those who hate or dislike a celebrity, ultimately posing potential harm to both fans and the celebrities involved.

[1.3] As technology reshapes social relations, fan communities' responses to scandals highlight the complexity of how fans, as a community, scrutinize and interact with technologies, and how those practices affect celebrities. This article may have begun as a set of methodological reflections; however, fans' navigation of the scandal raised significant questions about their perception of technologies. This presents an opportunity to rethink ethics in fandom research in an age where every click, every post, and every view is intertwined with AI and machine learning algorithms.

2. How it started: From a video scandal to a deepfake event

[2.1] One summer day in 2023, when we were doing our routine surfing and data collection for our research on the transnational fandom of a Thai GL television series, a scandal involving two Thai celebrities broke out on the Internet. A short video clip circulated on X (formerly called Twitter) and Weibo (a China-based social media platform) showing a man and a woman sharing a brief kiss on the lips. Filmed from a distance, likely from outside through the windows into someone's home, the nine-second footage appeared blurry, vaguely showing both parties' faces [1]. Speculation quickly arose that both parties in the video were Thai actors from BL and GL television series, queer media that fantasize about same-sex romance. The woman in the video was suspected to be Praew, a GL star who had a substantial international queer following due to the recent success of a GL series [2]. The video sparked immense outcry and frustration within Praew's fandom, which caught our attention.

[2.2] The Thai GL and BL productions, known as "series Y" or the Y industry, have become rising stars in the queer media industry, gaining phenomenal global popularity in the early 2020s (Baudinette 2023; Li and Pang 2024b). Actors and actresses in these series, such as

Praew, could amass millions of Instagram followers worldwide within just a few months. Having become two such followers of Thai GL and/or BL, we became acafans of series Y, interested in its transnational fandom and analyzing social media data after obtaining ethics approval from our respective Institutional Review Boards [3]. We have been observing posts and hashtags related to the Thai GL series, celebrities, and events, as well as fans' discussion in public posts on social media platforms such as X, Weibo, and Instagram. The discussion in this article is based on our reflections on the observation of social media interactions. Aware of the potential problems of lurking in digital fan spaces, we present our observations in aggregate form to ensure the anonymity of fans' discussions online (Markham 2012). In this capacity, we witnessed how the video scandal mentioned above developed into a debate and play involving deepfake.

[2.3] We witnessed how this video upset many in the Thai GL fandom on social media. This reaction is understandable because GL media, akin to BL, featured not only same-sex romance in television drama series but also the fantasy of real-person ship (RPS) between actresses who portray characters in romantic relationships on screen. That is, fans imagine (or expect) actresses to be real-life, off-screen lovers. Likewise, since GL media often portrays a utopian world for queer women (Li and Pang 2024b), fans are heavily invested in the queer fandom of the GL series in which Praew starred, as well as the RPS between Praew and her same-sex co-star [4]. Therefore, this video, although taken clandestinely and showing the two sharing a light, brief kiss on the lips, revealed shocking news to fans: that Praew might have a male lover in real life. Such a seemingly small act caught on camera could potentially shatter fans' queer fantasy because the footage suggests that Praew is not a lesbian in real life [5].

[2.4] Many fans, most of whom are queer women, felt betrayed and dismayed that the RPS was not "real"; that is, the on-screen couple were not lovers in real life. As soon as the video circulated online, some fans quickly questioned whether it was "real". This skepticism arises because spreading rumors about GL or BL stars engaging in heterosexual affairs has been a common tactic used by anti-fans to defame the stars and jeopardize their careers for being "dishonest" about their sexuality in real life. Among various speculations, the most notable was that the video was a deepfake, thus disqualifying it as credible evidence of Praew's sexuality. Such possibilities were seriously debated among fans until the official clarification statement was released three days later.

[2.5] Fans' frustration and denial of the authenticity of the video are not mere knee-jerk reactions to celebrity scandals; they manifest fans' deeper perception of technologies as well as their views of the Thai GL industry. In this particular scandal, fans' folk theories on social media reveal their perceptions and scepticism of how the media (or the machine) works, which informs their understanding of the incident (Ytre-Arne and Moe 2021; Xu et al. 2024).

3. Parodic deepfake as counter-information and digital play

[3.1] Within the three-day window between the video going viral and the release of the official statement from the artist management companies, we observed three folk theories that emerged on social media: (1) the video was "real", but the person was a look-alike of Praew; (2) fans trusted that the management company would be able to prove the video was fake; and (3) the video was a malicious deepfake, in which somebody else's face had been swapped with Praew's. Among these, the third theory—that the video was a deepfake—circulated quickly. Deepfakes, created using AI or machine-learning tools that combine, alter, and overlay images and video to produce realistic yet fake content, can be easily generated with open-source, user-friendly software and are designed for rapid dissemination on social media (Maras and Alexandrou 2019; Karnouskos 2020). Moreover, 2023 was a year where GenAI and deepfake were the talk of the town and became more accessible to ordinary people around the world. Researchers across disciplines have recently begun exploring the social implications of deepfakes, such as those related to legal contexts (Lau 2023; Maras and Alexandrou 2019), performance (Fletcher 2018), misinformation (Day 2019), and pornography (Popova 2020). Among these, a large-scale, cross-national quantitative research by Ahmed (2023, 324–25) revealed a telling result: deepfakes were more convincing than other forms of disinformation, and members of the public, despite being deceived, generally believed that they were less likely to be deceived by deepfakes than others. In the case Praew, when the video cannot be relied upon as evidence to verify or nullify information, further evidence is expected.

[3.2] Three days is a long time to wait for the official clarifications regarding the authenticity of the video; many fans could not wait. Some were eager to assess the "realness" of the video by conducting meticulous but relatively "low-tech" frame-by-frame dissection to look for traces of inconsistency and manipulation. Given the high technical threshold for deepfake

detection, which involves training data for detection models (Groh et al. 2024, Jacobsen 2024), it is not surprising that we were not able to locate any conclusive video forensic report shared by fans. That being said, research on folk theories of false information suggests (Koçer et al. 2022) that people tend to believe information from their own social circles. That could be why we kept encountering social media comments reiterating, "the video is deepfake". In other words, fans reproduced and perpetuated the belief that the video was a deepfake without seeing any professional reports of deepfake detection.

[3.3] Since detecting deepfakes was more difficult than expected, some fans looked for alternatives to illustrate how easy it is to create them. Approximately one day after the outbreak of the scandal, we witnessed the circulation of a parodic deepfake made from the original video. Parody has been a prevalent, if not central, element of participatory media culture (Boxman-Shabtai 2018). Making and circulating parodic deepfakes is a tactic that fans use to mock and undermine scandalous materials, thereby reducing the harm inflicted on both the fan community and the celebrity. The parodic deepfake we observed replaced the face of the woman, who was speculated to be Praew, with that of Shane, Praew's artist manager at the time. Shane was often seen at public events, so fans could easily find his footage and pictures on social media. The parodic deepfake showed a seamless face swap, closely resembling the facial expression in the original video. This parodic deepfake aimed to demonstrate that creating hyper-realistic deepfakes is not as difficult as people believe. We also noted that fans circulated and discussed this parodic deepfake primarily in a playful manner. This was partially because Shane was well-liked by fans for his affability, which made the parodic deepfake "relatable" and allowed for queer pleasure in shipping him with the BL actor involved in the video. Additionally, the parodic deepfake served as comic relief for the fandom during a tough time following the scandal. Fans who saw the parodic deepfake jokingly expressed on social media that Shane and the BL actor "were a perfect match," and they would be happy to fantasize about them as lovers. With such a parodic video, the "realness" of the original video was also undermined.

[3.4] This form of AI-generated parodic image is unlike fans' other longstanding creative practices. Media users can quickly identify memes or GIFs that use celebrities' or public figures' faces for entertainment due to their recognizable style (Popova 2020). In contrast, the hyper-realistic parodic deepfakes made by fans risk disseminating believable disinformation,

even if it was not the fans' intention to do so or if they did so out of goodwill. Here, the person whose face was swapped was Shane, an artist manager known to the fandom but not entirely a public figure or celebrity. Such playful yet believable and spreadable disinformation could have impacted him personally. Indeed, the parodic deepfake was not intended for circulation beyond the fan community. Still, it is difficult to ensure that they remain confined to specific Internet spaces, given the ever-changing social media algorithms. Once these parodic deepfakes are circulated, they will be unerasable and retrievable online, subjecting all the individuals involved to doxing and cyberbullying. This is evident from what we have observed since the summer of 2023: the parodic deepfake images in this scandal were occasionally used by anti-fans to make fun of Shane and defame or attack Praew.

[3.5] More importantly, fans' tactics of countering difficult information or perceived deepfakes by creating parodic deepfakes might lead to unintended consequences for all parties involved or perceived to be involved. To be clear, we deplore paparazzi and stalkers and do not endorse the policing of celebrity sexuality, but what sparked a debate between us about this incident is the ethics of representation in fandom research. Three days later, on their official X account, Praew's management posted a pre-recorded video in which Praew apologized to the public and admitted that the video was taken without her consent. Without a doubt, any video taken without the consent of the parties involved should not be circulated or disseminated in any form—including academic research and publications. Nevertheless, how about the parodic deepfakes made by fans in which Shane's face was swapped? For pornographic deepfakes, where women (celebrities or ordinary people) are usually the victims, it is universally agreed that such images should not be disseminated by any means (Maddocks 2020; Story and Jenkins 2023). However, the parodic deepfake concerned here is more ambivalent. How should researchers study and represent the deepfaked body in academic research, and should they do so? Suppose we accept that digital images are subject to acceptable manipulation, such as applying filters, adjusting colors, or even removing objects with the help of AI-assisted functions in software; why should deepfakes be treated differently?

[3.6] We were caught in a dilemma. In research contexts, a digital image of the body includes both digital data and the material body, intersecting with complex ethical issues such as identifiability, confidentiality, anonymity, and publicness and privacy (Warfield et al. 2019, 2069). What distinguishes the deepfaked body from other digital images of the body is that the

deepfaked body involves the bodies of multiple individuals, typically consisting of the body from the source video and the face from another body that has been swapped. To be more concise, the deepfake in our case includes Shane's face and Praew's body. This adds complexity in terms of both the ethics of engagement and the ethics of representation. In terms of consent for using fans' content, it would be challenging, if not impossible, to trace the source of the parodic deepfake and obtain permission from the creator and all the involved individuals. Deepfakes as digital artifacts complicate the conventional digital-material nexus of images collected from research participants, such as selfies and personal portraits. In terms of representation, how shall we analyze the deepfake images and disseminate them as research findings?

4. Rethinking the ethics and politics of care in fandom research

[4.1] When researching fans' engagement with deepfake, at least two levels of representations complicate each other. First, deepfake technology manipulates representations based on the already available visual and/or audio materials. Second, researchers writing about deepfakes and fans' responses in academic publications are also engaging in a form of representation. Taking this further, such scholarly representation potentially contributes to machine learning datasets. While we recognize the academic relevance and significance of discussing fans' perceptions and the use of AI in their responses, we are also acutely aware of the potential ethical issues related to representing this case.

[4.2] To ensure that fans were not re-traumatized as a result of our writing about the scandal, we turn to the literature on research ethics in fan studies. "Fans first", the foundation of ethical practices in fandom research, was a direct response to research that pathologizes fandom and disrespects fans' interests (Hellekson and Busse 2009). Good practices include not including direct URLs to fans' pages or fanfiction in publications and obtaining consent when using fans' work for research. Dym and Fiesler (2020) also encourage researchers to understand fandom's community and privacy norms. When using online fandom data, researchers should ensure proper consent from and acknowledgement to the fans concerned while obscuring identifiable personal information. Importantly, they should return research to the fandom with research with goodwill, not harm (Dym and Fiesler 2020). Similarly, The Association of Internet Researchers (2012) ethics working committee upholds the primary principle of "do no harm". Because fans have constantly been pushing legal and moral boundaries through their

transformative works, it is particularly important to protect fan communities by not imposing harm and risks of harm through research, whether physical, psychological, financial, or through damage to reputation and intellectual property rights. Examples of "harm" in the context of fandom research include outing LGBTQ fans, the legal consequences of derivative work, and subjecting individuals or communities to online and offline harassment (Busse 2017). Nonetheless, "harm" in research should be defined contextually (Markham and Buchanan 2012). In the celebrity scandal discussed above, we should not only consider and respect fans' well-being, but also that of Shane and Praew, even though Praew is a celebrity.

[4.3] AI-generated content can unexpectedly harm public figures and the public, whether intended as parody, satire, or disinformation. Glick (2023) highlighted incidents like a deepfake of Gabonese President Ali Bongo's 2019 New Year's video, which led to an unsuccessful coup, and deepfakes of Ukrainian President Volodymyr Zelensky used by Russians to demoralize Ukrainian soldiers. In entertainment, celebrities like Nicolas Cage and Bruce Willis have faced reputational harm from AI-generated content. Cage's image has been used in numerous memes, reinforcing his reputation as a "bad actor," while Willis's face was used in a Russian ad without consent, raising political sensitivity during US sanctions on Russia (McGowan 2017). These cases demonstrate the undeniable harm caused by deepfakes.

[4.4] Without anyone being held accountable for the highly believable deepfake and its accompanying disinformation, celebrities' fear of losing control over their images and bodies is intensified, especially given the viral nature of social media. The long-term implications of deepfakes are unpredictable due to the lack of transparency in data usage for machine learning across various AI models (Cao and Yousefzadeh 2023). Unlike Busse's (2017) focus on text-based fandom harms, such as fanfiction, Deller (2018) argues for considering a broader range of agents, including individuals, communities, media producers, public figures, and even physical objects. Given the networked nature of machine learning, we find Deller's approach more applicable to understanding harm in fandom research when AI is involved. In particular, in fandom, where the subject of fandom is an embodied person, harm could be inflicted not only on fans but also on public figures or celebrities.

[4.5] In the case of Praew discussed here, if we are to represent the parodic deepfake image in any visual form, it would be ideal—but almost impossible,—to have obtained permission from the fan who created this deepfake image, as well as from all the individuals concerned, namely, Praew, Shane and the BL actor involved. Even if we have the consent to do so for academic purposes, it is unnecessary to reproduce and display the images because a screenshot of the parodic deepfake would not add anything to our discussion, except perhaps a laugh at academic conferences, or serve as “supplementary” information in academic publications, since we can textually describe it along with other contextual information. It is unwise to provoke such a “joke” at the expense of potentially exacerbating the trauma experienced by the already affected fandom and contributing to the machine learning datasets with disinformation. As explained above, online records of the leaked video and deepfake visuals have already been used for cyberbullying and doxing.

[4.6] This line of thinking aligns with the reflexive position of acafandom—putting a hyphen between aca(demic) and fandom—which is an implementation of “fans first” and “do no harm” in practice. It implies a certain level of care—the care for the fan community that a researcher identifies with or is part of. However, whether occupying such a position produces good scholarship and ethical practice has been much debated (Brooker, Duffett, and Hellekson 2017; Hills 2012, 2002). This is because acafans not only have the power to represent fandom through their research but also silence the questionable aspects within fandom (Brooker, Duffett, and Hellekson 2017; Busse and Hellekson 2012). Meanwhile, both fandom and social media are fragmented spaces, especially within the algorithmic culture of social media platforms, which often reinforces filter bubbles and echo chambers. The position of acafan does not guarantee understanding all the subgroups within fandom. In the context of this scandal, our acafan position presented challenges. Emotionally, we were upset that the video was taken without Praew’s consent, and we understood why some queer fans were distressed. However, we were deeply concerned when fans trolled and attacked Praew for “betraying” them, as fans jeopardizing actors’ mental well-being is not uncommon in the Thai Y industry. This prompts us to consider the issue of care when fandom intersects with technology—what does it mean for the notion of “care” in the context of fandom research? Is “fans first” as an overarching ethical framework sufficient in the age of AI?

5. Towards a critically speculative ethics of care in fandom studies

[5.1] Given the highly commercialized and industrialized structure of the contemporary entertainment industry, celebrities are often treated as signs, texts, or signifiers. There is also a common assumption that celebrities are protected by capital and resources that make them invincible, physically and emotionally. We observe a contradictory development where celebrities increasingly emphasize their humanness, such as vulnerability, to appear more “authentic” and relatable. At the same time, the commodification of their bodies and personas simultaneously makes them prone to being perceived as objects of entertainment, subject to unreasonable demands in the name of “professionalism”, which dictates that actors should fulfil fans’ fantasy 24/7, even in their private time and space. Similarly, in fandom research, we tend to celebrate fans’ agency and participatory practices instead of addressing the experiences of celebrities and other individuals involved or affected. This focus is understandable, as fans are the primary subject matter in fandom studies, and the entertainment industry often benefits from the fruits of fan labor, such as fame and popularity (Stanfill and Condis 2014). However, it is crucial to recognize that celebrities and other concerned individuals can be susceptible to various forms and degrees of harm, and that “fandom” consists not only of fans but other human subjects (celebrities and industry workers) and human-influenced technologies, such as digital platform algorithms and AI.

[5.2] After all, context matters. Wilfully ignoring the significance of contexts in shaping perceptions of celebrity scandals and fans’ responses risks replicating the coloniality present in academic knowledge production, which has been extensively critiqued in transnational feminist scholarship (Falcón 2016, Zerbe Enns et al. 2021). We asked ourselves: Is the video considered sensitive within its specific context, including the national/regional context and the context of a particular fandom? Compared to most Euro-American societies, where sexuality can be relatively openly discussed and celebrities can speak out for themselves, public displays of affection remain a taboo in many Asian societies. Even though Praew’s leaked video was non-consensual, the footage has entered the (semi-)public domain on the Internet. We often see similar scandals involving female celebrities’ sexuality or intimate lives in Asia, resulting in public apologies for disappointing their fans rather than defending their own rights [6]. This is why we debated how best to implement the speculative ethics of care in our writing while contributing to the emerging scholarly discussion on the broader impact of AI on fandom and society.

[5.3] We understand that researchers should commit to upholding academic integrity to provide sufficient contextual information for readers to understand the complexities of the issue and ground a critical discussion in our methodological reflections. This is particularly important because non-Western studies often necessitate more contextual details than their Western counterparts. Yet, this additional context creates a paradoxical challenge: it is difficult to include within the limited space of academic journals or presentations. Such details may appear overly descriptive compared to Western-focused work, despite their critical role in addressing the unfamiliarity of the non-Western setting.

[5.4] At the same time, we should minimize, if not eliminate, all possible risks of harm to the best of our ability. Indeed, readers may still be able to infer more details of the event from the academic footprint of the author; however, using pseudonyms and masking or removing direct identifiers are measures we could adopt as long as they do not conflict with the public interest. We argue that such ethics of care is particularly relevant in this age of AI, as digital creativity and artifacts made by both fans and anti-fans contribute to the existing and potential datasets in machine-learning models to train future AIs. Thus, the fan community and fandom research can contribute to such processes. For example, we have witnessed Praew being a continuous target of physical and online harassment since the airing of the GL series in which she starred. Hence, participatory practices in fandom not only shape the semi-public fan communities but also influence social media algorithms and the machine learning dataset of AI.

[5.5] Building upon our observations on fans' entanglement with AI in response to Praew's scandal, we propose a speculative ethics of care for fandom studies researchers investigating fans' interactions with AI-generated images. The word speculative here does not mean reckless claims but rather intellectual precaution that is imaginative, pre-emptive, and proactive. The speculative, as argued by feminist technoscience philosopher Maria Puig de la Bellacasa, indeed "connects to a feminist tradition for which this mode of thought about the possible is about provoking political and ethical imagination in the present" (2017, 7). Recent examples have shown that public figures can be victims of AI-generated visuals. Still, given the black-boxing of AI and machine learning, scholars in the humanities and social sciences can only critically speculate, with care, on the potential harms of technologies in relation to

fan practices, such as parodic deepfakes and other AI-generated creativity (Amoore et al. 2023; Birrer and Just 2024).

[5.6] "Care" has been proposed by feminist philosophers as an alternative approach to moral philosophy (Keller and Kittay 2017). Care, in the broadest sense, refers to "a species of activity that includes everything we do to maintain, contain, and repair our 'world' so that we can live in it as well as possible" (Fisher and Tronto 1990, 40). The feminist ethics of care is, therefore, "a critical approach that seeks to understand the necessity of care to well-being, to understanding marginalization and identifying responsibility to remedy social injustices" (Brannelly and Barnes 2022, 6). While the feminist ethics of care have influenced a range of academic disciplines, such as healthcare, economics, and international relations (see Keller and Kittay 2017, 545), Puig de la Bellacasa (2017) contends that the definition of care is also a speculative project in its own right. The principle of "fans first" prioritizes the interest and well-being of fans, but when considering the intersectional and complex relationship between human subjects, community, and technologies in fandom, we have to think beyond "fans first". Drawing upon Puig de la Bellacasa (2017), we articulate three guiding principles of the speculative ethics of care for fandom researchers that call for a rethinking of fandom research ethics beyond "fans first":

[5.7] Thinking with fans

This refers to a relational way of thinking, which involves a commitment to a collective of knowledge-making (Puig de la Bellacasa 2017). This is closely related to "fans first", as thinking with fans focuses on empathizing with them and staying attuned to the sociocultural trajectories of their desires and feelings while still addressing both the scandal and the potential harm it might bring to fans. This also means respecting fans' wishes to protect celebrities and others involved from further Internet doxing and cyberbullying. One might be concerned that thinking with fans means foregoing academic rigor and integrity, but academic research is not merely a journalistic aggregation of fans' voices. Thinking with fans involves a thoughtful process of showing care to the community we research. In our case, we continually ask ourselves questions such as "How can we write about this scandal and fans' responses without triggering fans?"

[5.8] Thinking for fandom

Taking into consideration the feelings of fans, acafan can also contribute to fandom by offering alternative perspectives. Care often implies love and kindness, and so does the idealized perception of fandom as a loving, inclusive community that transcends linguistic and geographical barriers. However, fandom is also a site of conflict and disagreement. Considering fandom as an interdependent, relational web of relationships, it is increasingly filled with tension, especially on social media platforms where algorithms place us in filter bubbles and echo chambers. Researching fandom with care not only means thinking with and empathizing with fans but also being aware of tensions and dissent, hierarchies and power dynamics, as well as how existing relationality contributes to new forms of disconnection in relationships. Building care while recognizing divergent positions and offering new insights that may differ from them can be challenging, but we should not be shy away from this responsibility.

[5.9] Thinking beyond fans and fandom

Fandom research should care for fans and beyond. One must respect everyone involved in the analyzed case, not only fans but also their subject(s)-of-fandom and other public or semi-public figures, and see them as embodied human beings with feelings and dignity. Puig de la Bellacasa (2017, 85) warns us of the danger of confusing ourselves as researchers with spokespersons who use marginalized others (such as fans) as arguments. In the case of leaked video and parodic we discussed above, the "others" include celebrities and the manager. As researchers, we must be aware of the multilayered representations at play, which complicate research of this kind. Studying fan engagement with deepfakes requires grappling with the represented nature of deepfakes themselves as well as the scholarly writing about them, which can feed into machine learning datasets and shape future knowledge. Hence, thinking beyond fans and fandom involves acknowledging what "fandom" means, as participatory culture has been transformed and reshaped by technology, as indicated by HCMI (Li and Pang 2024a). In participating in the fandom, we are also participating in the world of AI. Even though fans remain the core focus of inquiry in this field, researchers should shoulder the responsibility of speculating and minimizing the possible risks of harm to all concerned subjects.

[5.10] In addition to the basic ethical considerations and measures, such as obtaining approval from the institutional research ethics board and anonymizing the fans involved, we have taken a step further. As a praxis of the critically speculative ethics of care we proposed, we decided to pseudonymize the celebrities involved and not provide any screenshots or URLs to the images. It is critically speculative because no one knows precisely the extent and timescale of harm those images will cause in the age of machine learning and GenAI. Suppose both fans and celebrities consider the scandal traumatizing; in that case, there is no reason to exacerbate the situation by reproducing those images in the name of academic knowledge production while relinquishing ethics of care.

6. Conclusion

[6.1] Fans have always been the first among media users to experiment with new technologies. The current AI revolution is no exception. In parallel, fandom scholars also need to start addressing ethical issues when AI is involved in fans' parodic play and creative practices. From the puzzling question "Is it deepfake?" to the assertion "It is deepfake!", we have observed fans' approaching celebrity scandals by scrutinizing the authenticity of the scandalous video or using AI/face swap to counter such difficult information through playful disinformation. In other words, considering fandom as Human-Community-Machine Interactions (HCMI), when individual fans ("human") and the broader fandom ("community") were distressed by the video scandal, technology ("machine") becomes the resort to verify or discredit the scandalous video. These practices, such as video dissection and the creation of parodic deepfakes, are further amplified on social media ("machine") and engaged by fans ("community"). On the one hand, parodic deepfakes showcase participatory creativity. On the other hand, they pose potential risks, such as spreading disinformation and contributing it to machine learning datasets, which could lead to unforeseeable consequences. In short, the interactions between fan communities and digital technologies play a significant role in the collective process of socio-technological meaning-making. With this awareness in mind, we feel like we had opened Pandora's box.

[6.2] Reflecting on the challenges of researching this deepfake event, we raise questions about representation and harm in fandom research within a digital ecology where AI and machine learning transform the very meaning of fandom as participatory culture. The proposed

speculative ethics of care expands the “fans first” principle to include not only fans but also celebrities, industry workers, other human agents, and HCMI. With the emergence of deepfakes, fans’ collective responses through digital play offer a significant area for exploration. Researchers must address the vulnerability of both fans and celebrities, the tensions within fandom, and the complex relationships among fans, celebrities, industry workers, and other social agents. We hope this article serves as a springboard for further discussions on ethics and care in fandom research concerning AI-generated playful artifacts, (dis)information, and the relevance of AI governance to fandom.

[6.3] Ethical dilemmas are not binary—research or not to research, write or not to write. In a digital landscape increasingly shaped by GenAI, data is both a source and vehicle of power (Hasselbalch 2021). This article has only begun to explore the ethical implications of researching fans’ engagement with AI-generated content, and we acknowledge the constraints of conclusions drawn from reflections on a single case. While we are still waiting to see AI’s more profound societal impact, our insights suggest avenues for future research on AI and fandom. For example, what if the parodic deepfake discussed earlier is used for training future AI about Praew, GL, or queer fandom? Recent developments, such as the introduction of AI-generated features on the wiki hosting service Fandom (David 2024) and AI-generated fanart and fanfic (Lamerichs 2018; Mussies 2023), pose uneasy questions for fans: What if the products of fan labor, such as discussions and artwork, are used to train AI models that inform future applications? How might this affect fan communities, celebrities, and media producers? Further research is needed to examine how fans contribute to AI machine learning, the implications for fan communities, and the ethics of fandom research. Researchers should be aware of the potential harm caused by AI, deepfake, and online disinformation by conducting reflexive studies that uphold empathy and care.

[6.4] This deepfake event in the Thai GL fandom opens Pandora’s box, prompting critical methodological reflections on research ethics despite the challenges it poses. The key to this critically speculative ethics is ensuring we do not harm fans and others affected by our knowledge production. The three proposed principles—thinking with fans, thinking for fandom, and thinking beyond fans and fandom—serve as a starting point for re-examining fandom and participatory culture in the age of GenAI. When thinking with fans, researchers should envision themselves as the subjects in a deepfake scandal and critically reflect on

whether they would be comfortable being treated and represented in that manner. In thinking for fandom, researchers, especially acafans, must consider the existing relationships that can lead to internal schisms, tensions, or disconnection, and navigate care within that context. Additionally, thinking beyond fans and fandom involves critically evaluating the potential impact of their findings when they are used to train AI models, which may affect fans and celebrities in ways that are not immediately evident. Understanding fandom in the HCMI framework (Li and Pang 2024a) helps us better situate fannish practices within an evolving digital landscape. In this way, fandom research remains a reflexive project of co-constructing knowledge with fans—an endeavor that is not only intellectually rigorous but also ethical, healing, and empowering.

7. Acknowledgements

[7.1] We would like to thank Carolyn Pedwell and the anonymous reviewers for their comments on the earlier version of the article.

8. Note

[1] We decided not to include a direct link to the video as it was maliciously disseminated. We will discuss our decision further in the rest of the article.

[2] Praew is a pseudonym. The same applies to the name Shane later in this article. In later sections, we will discuss why pseudonyms are used in this article.

[3] Transnational fandom here refers to non-Thai fans.

[4] We use “queer” as an umbrella term to encompass a wide range of queer responses, including all non-straight expressions, identities, and desires (Doty 1993). Thus, our use of “queer fandom” refers not only to non-straight fans but also to straight fans who find pleasure and resonance in consuming queer media.

[5] Bi-erasure was observed in the fandom as there was only little discussion on the possibility that Praew might be bisexual or pansexual. The methodological basis of our observation is elaborated later in the article.

[6] For example, Japanese pop star Minami Minegishi of AKB48 shaved her head and apologized for spending a night with her boyfriend, while K-pop star Karina of the girl group aespa publicly apologized for having a boyfriend (McCurry 2013, Ng 2024).

9. Works Cited

- Ahmed, Saifuddin. 2023. "Examining Public Perception and Cognitive Biases in the Presumed Influence of Deepfakes Threat: Empirical Evidence of Third Person Perception from Three Studies." *Asian Journal of Communication* 33, no. 3: 308–31. <https://doi.org/10.1080/01292986.2023.2194886>.
- Amoore, Louise, Alexander Campolo, Benjamin Jacobsen, and Ludovico Rella. 2023. "Machine Learning, Meaning Making: On Reading Computer Science Texts." *Big Data & Society* 10, no. 1: 1–12. <https://doi.org/10.1177/20539517231166887>.
- Baudinette, Thomas. 2023. *Boys Love Media in Thailand: Celebrity, Fans, and Transnational Asian Queer Popular Culture*. London: Bloomsbury Publishing.
- Birrer, Alena and Just, Natascha. 2024. "What we know and don't know about deepfakes: An investigation into the state of the research and regulatory landscape" *New Media & Society*, Epub ahead of print. <https://doi.org/10.1177/14614448241253138>
- Boxman-Shabtai, Lilian. 2019. "The practice of parodying: YouTube as a hybrid field of cultural production." *Media, Culture & Society*, 41, no.1: 3–20. <https://doi.org/10.1177/0163443718772180>
- Brannelly, Tula and Barnes, Marian. 2022. *Researching with Care: Applying Feminist Care Ethics to Research Practice*. Bistol: Policy Press.
- Brooker, Will, Mark Duffett, and Karen Hellekson. 2017. "Fannish Identities and Scholarly Responsibilities: A Conversation." In *The Routledge Companion to Media Fandom*, edited by Melissa A. Click and Suzanne Scott, 63–74. New York: Routledge.
- Browne, Jude, Stephen Cave, Eleanor Drage, and Kerry McInerney. 2023. *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*. Oxford University Press.
- Busse, Katrina, and Karen Hellekson. 2012. "Identity, Ethics, and Fan Privacy." In *Fan Culture: Theory/Practice*, edited by Katherine Larsen and Lynn Zubernis. Newcastle upon Tyne: Cambridge Scholars Publishing.

- Busse, Kristina. 2017. "The Ethics of Studying Online Fandom." In *The Routledge Companion to Media Fandom*, edited by Melissa A. Click and Suzanne Scott, 9–17. New York: Routledge.
- Cao, Xuenan, and Roozbeh Yousefzadeh. 2023. "Extrapolation and AI Transparency: Why Machine Learning Models Should Reveal When They Make Decisions beyond Their Training." *Big Data & Society* 10, no. 1: 1–5.
<https://doi.org/10.1177/20539517231169731>.
- David, Emilia. 2024. "Fan wiki hosting site Fandom rolls out controversial AI features". Available at: <https://www.theverge.com/2024/2/12/24070287/fandom-wiki-generative-ai-features-quick-answers> (accessed 2 August 2024).
- Day, Charles. 2019. "The Future of Misinformation." *Computing in Science & Engineering* 21, no. 1: 108. <https://doi.org/10.1109/MCSE.2018.2874117>.
- Deller, Ruth A. 2018. "Ethics in Fan Studies Research." In *A Companion to Media Fandom and Fan Studies*, edited by Paul Booth, 123–40. Campbell, CA: John Wiley & Sons.
- Doty, Alexander. 1993. *Making Things Perfectly Queer: Interpreting Mass Culture*. Minneapolis: University of Minnesota Press.
- Dym, Brianna, and Casey Fiesler. 2020. "Ethical and Privacy Considerations for Research Using Online Fandom Data." *Transformative Works and Cultures*, no. 33. <https://doi.org/10.3983/twc.2020.1733>
- Falcón, Sylvanna M. 2016. "Transnational Feminism as a Paradigm for Decolonizing the Practice of Research: Identifying Feminist Principles and Methodology Criteria for US-Based Scholars" *Frontiers: A Journal of Women Studies* 37, no. 1: 174–194.
- Fisher, Berenice, and Joan Tronto. 1990. "Toward a Feminist Theory of Caring." In *Circles of Care: Work and Identity in Women's Lives*, edited by Emily K. Abel and Margaret K. Nelson, 35–62. Albany, NY: State University of New York Press.
- Fletcher, John. 2018. "Deepfakes, Artificial Intelligence, and Some Kind of Dystopia: The New Faces of Online Post-Fact Performance". *Theatre Journal* 70, no. 4: 455–71.
<https://doi.org/10.1353/tj.2018.0097>.
- Glick, Joshua. 2023. Deepfake Satire and the Possibilities of Synthetic Media. *Afterimage* 50, no. 3:: 81–107.

- Groh, Matthew, Aruna Sankaranarayanan, Nikhil Singh, Dong Young Kim, Andrew Lippman, and Rosalind Picard. 2024. "Human Detection of Political Speech Deepfakes Across Transcripts, Audio, and Video." arXiv.
<https://doi.org/10.48550/arXiv.2202.12883>.
- Hasselbalch, Gry. 2021. *Data Ethics of Power: A Human Approach in the Big Data and AI Era*. Cheltenham, UK Northampton, MA, USA: Edward Elgar Publishing Ltd.
- Hellekson, Karen, and Kristina Busse. 2009. "Fan Privacy and TWC's Editorial Philosophy." *Organization for Transformative Works*. December 5, 2009.
<https://www.transformativeworks.org/twac/twc-citation/>.
- Hills, Matt. 2012. "'Proper Distance' in the Ethical Positioning of Scholar-Fandoms: Between Academics' and Fans' Moral Economies?" In *Fan Culture: Theory/Practice*, edited by Katherine Larsen and Lynn Zubernis, 14–37. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Jacobsen, Benjamin N. 2024. "Deepfakes and the promise of algorithmic detectability." *European Journal of Cultural Studies*. Epub ahead of print. <https://doi.org/10.1177/13675494241240028>
- Karnouskos, Stamatis. 2020. "Artificial Intelligence in Digital Media: The Era of Deepfakes." *IEEE Transactions on Technology and Society* 1, no. 3: 138–47.
<https://doi.org/10.1109/TTS.2020.3001312>.
- Keller, Jean and Kittay, Eva Feder. 2017. "Feminist Ethics of Care." In *The Routledge Companion to Feminist Philosophy*, edited by Ann Garry, Serene J. Khader, and Alison Stone, 540–555. New York: Routledge.
- Koçer, Suncem, Bahadır Öz, Gülten Okçuoğlu, and Fezal Tapramaz. 2022. "Folk Theories of False Information: A Mixed-Methods Study in the Context of Covid-19 in Turkey." *New Media & Society*. <https://doi.org/10.1177/14614448221142310>.
- Lamerichs, Nicolle. 2018. "The Next Wave in Participatory Culture: Mixing Human and Nonhuman Entities in Creative Practices and Fandom." *Transformative Works and Cultures* 28. <https://doi.org/10.3983/twc.2018.1501>.
- Lau, Timothy. 2023. "'Pics or It Didn't Happen' and 'Show Me the Receipts': A Folk Evidentiary Rule." *Vanderbilt Law Review* 76, no. 6: 1681–1707.
<https://heinonline.org/HOL/P?h=hein.journals/vanlr76&i=1713>.

- Li, Eva Cheuk-Yin, and Ka-Wei Pang. 2024a. "Fandom Meets Artificial Intelligence: Rethinking Participatory Culture as Human–Community–Machine Interactions." *European Journal of Cultural Studies* 27, no. 4: 778–787.
<https://doi.org/10.1177/13675494241236146>
- Li, Eva Cheuk-Yin, and Ka-Wei Pang. 2024b. "Girls' Love Media Industry in Southeast Asia: A Mere Replica of Boys' Love Media or New Routes for Queer Representation and Allyship?" Conference paper presented at *Media Industries*, King's College London, UK, 17 April.
- Maddocks, Sophie. 2020. "'A Deepfake Porn Plot Intended to Silence Me': Exploring Continuities between Pornographic and 'Political' Deep Fakes'." *Porn Studies* 7, no. 4: 415–23. <https://doi.org/10.1080/23268743.2020.1757499>.
- Maras, Marie-Helen, and Alex Alexandrou. 2019. "Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos." *The International Journal of Evidence & Proof* 23, no. 3: 255–62.
<https://doi.org/10.1177/1365712718807226>.
- Markham, Annette. 2012. "Fabrication as Ethical Practice: Qualitative inquiry in ambiguous Internet contexts" *Information, Communication & Society* 15, no.3: 334–353.
- Markham, Annette, and Elizabeth Buchanan. 2012. "Recommendations from the AoIR Ethics Working Committee (Version 2.0)." Association of Internet Researchers.
<https://www.aoir.org/reports/ethics2.pdf>.
- McGowan, David. 2017. "Nicolas Cage – good or bad? Stardom, performance, and memes in the age of the Internet" *Celebrity Studies* 8, no.2: 209–227.
- McCurry, Justin. 2013. "Japanese pop star shaves head in apology – for night with boyfriend." *The Guardian*, 1 February. Available at:
<https://www.theguardian.com/world/2013/feb/01/japanese-pop-star-apology-boyfriend> (accessed 1 August 2024).
- Mussies, Martine. 2023. "Artificial Intelligence and the Production of Fan Art." *Transformative Works and Cultures* 40. <https://doi.org/10.3983/twc.2023.2431>.
- Ng, Karry. 2024. "K-pop star Karina apologises after relationship goes public." *BBC News*. 6 March. Available at: <https://www.bbc.com/news/world-asia-68487232> (accessed 1 August 2024).

- Popova, Milena. 2020. "Reading out of Context: Pornographic Deepfakes, Celebrity and Intimacy." *Porn Studies* 7, no. 4: 367–81.
<https://doi.org/10.1080/23268743.2019.1675090>.
- Puig de la Bellacasa, Maria. 2017. *Matters of Care: Speculative Ethics in More than Human Worlds*. Minneapolis, MN: University of Minnesota Press.
- Stanfill, Mel and Condis, Megan. 2014. "Fandom and/as labor". *Transformative Works and Cultures* 15. <https://doi.org/10.3983/twc.2014.0593>.
- Story, Daniel, and Ryan Jenkins. 2023. "Deepfake Pornography and the Ethics of Non-Veridical Representations." *Philosophy & Technology* 36, no. 56: 1-22.
<https://doi.org/10.1007/s13347-023-00657-0>.
- The Association of Internet Researchers (AoIR). 2012. "Ethical Decision-Making and Internet Research". <https://aoir.org/reports/ethics2.pdf>.
- Warfield, Katie, Jamie Hoholuk, Blythe Vincent, and Aline Dias Camargo. 2019. "Pics, Dicks, Tits, and Tats: Negotiating Ethics Working with Images of Bodies in Social Media Research." *New Media & Society* 21, no. 9: 2068–86.
<https://doi.org/10.1177/1461444819837715>.
- Xu, Liying, Yuyan Zhang, Feng Yu, Xiaojun Ding, and Jiahua Wu. 2024. "Folk Beliefs of Artificial Intelligence and Robots." *International Journal of Social Robotics* 16: 429–446. <https://doi.org/10.1007/s12369-024-01097-2>.
- Ytre-Arne, Brita, and Hallvard Moe. 2021. "Folk Theories of Algorithms: Understanding Digital Irritation." *Media, Culture & Society* 43, no. 5: 807–24.
<https://doi.org/10.1177/0163443720972314>.
- Zerbe Enns, Carolyn, Díaz, Lilian Comas and Bryant-Davis, Thema. 2021. "Transnational Feminist Theory and Practice: An Introduction." *Women & Therapy* 44, no. 1–2: 11–26. <https://doi.org/10.1080/02703149.2020.1774997>