A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combating Adolescent Sexting Behaviors

Muhammad Uzair Tariq(✉), Afsaneh Razi(✉), Karla Badillo-Urquiola(✉), and Pamela Wisniewski(✉)

University of Central Florida, Orlando, FL, USA
uzairtq@gmail.com, {afsaneh.razi, kcurquiola10}@knights.ucf.edu, pamwis@ucf.edu

Abstract. We present a comprehensive literature review of algorithmic or computational approaches to detect nudity and/or skin that could be used as a means of preventing adolescent sexting behaviors. We identified 45 peer-reviewed articles that summarize the state-of-the-art in this field to show research gaps and opportunities for future research. We found several important gaps in the literature. For instance, most of the work related to the detection of nudity and/or skin has been done at the software level only; and while numerous algorithms exist in this space, they all operate on already digitized images. Therefore, researchers should consider addressing nudity detection at the hardware level to prevent digitization of these images before they cause harm. In addition, most of the literature we reviewed focused on the computational aspects of detection without further exploring what interventions may be appropriate once detection has occurred. Hence, they do not meaningfully address risk mitigation strategies that would be effective for tackling the problem of adolescent sexting behaviors. Therefore, computational researchers who focus on nudity and/or skin detection need to engage with Human-Computer Interaction (HCI) researchers to determine how to translate their research into actionable solutions for adolescent online safety.

Keywords: Computer vision · Nudity detection · Skin detection · Sexting · Adolescent online safety

1 Introduction

According to a recent survey by Pew Research Center, nearly 95% of adolescents have access to the internet and almost half (45%) of them are constantly online [1]. The internet provides great advantages, such as information sharing. Yet, it also introduces numerous problems dealing with accessing and sharing of explicit content, including sexting, cyberbullying, and child pornography [2]. Prolific sharing, combined with the permanence of digitally captured nudity, is particularly problematic for minors. While the dissemination of child pornography is a crime punishable by law [3], such

M. Kurosu (Ed.): HCII 2019, LNCS 11568, pp. 90–108, 2019.
https://doi.org/10.1007/978-3-030-22636-7_6

momentary mistakes could also cause physical harm and prolonged psychological problems for adolescents, including sexual predation, emotional trauma, cyberbullying, and even suicidal behaviors [4].

It is estimated that about 15% of teens on Snapchat report having received sexually explicit photos. In addition, 4% of cellphone-owning teens, ages 12–17, report having sent sexually suggestive, nude, or nearly nude images of themselves to someone else via text messaging [5]. Sexting behaviors make adolescents vulnerable to a number of offline risks, such as bullying [2] and sexual predation [6]. These types of teen sexting behaviors can be perpetuated by mobile technologies and by several direct messaging applications available on smartphone devices, such as Kik, Snapchat, and AskFM [7]. Unfortunately, such activities often fall under the jurisdiction of child pornography laws. Child pornography is illegal, and the federal law states the following: "A picture of a naked child may constitute illegal child pornography if it is sufficiently sexually suggestive. Additionally, the age of consent for sexual activity in a given state is irrelevant; any depiction of a minor under 18 years of age engaging in sexually explicit conduct is illegal" [3]. So, while sexting behaviors may seem innocent and exploratory to teens, in reality, they can have severe negative consequences.

Solutions have been proposed to approach the problem of adolescent sexting within the overlapping boundaries of various disciplines. For example, this topic has been studied within computer science [8], psychology [9], communications [10], and digital forensics [2]. Overall, very few actionable solutions have been proposed, but the core focus of computer science research has been on the accurate detection of nudity and/or skin as a means of prevention. Computational researchers have used different methods of machine learning, which we synthesize in this paper. Specifically, we investigated the following research questions in topic of nudity and/or skin detection:

- **RQ1:** *What characteristics are common among existing research on nudity and/or skin dominance detection?*
- **RQ2:** *What algorithmic approaches have been used in this literature?*
- **RQ3:** *What research or user studies have been conducted to translate this research into real-world interventions for adolescent online safety?*

We present a comprehensive literature review of algorithmic and computational approaches to detect nudity and/or skin that could be used as a means of preventing teen sexting behaviors to answer these research questions. Our main findings are that most research related to nudity detection has been studied after such images have been captured (RQ1), so there is a need to prevent capturing such images at first place. General solutions for nudity detection should be applicable on mobile platforms. Also, visual features in images should consider detecting the age range of the subject to protect adolescents and minorities from engaging in this behavior (RQ2). Finally, more risk mitigation solutions should be proposed in addition to risk detection and user-centered design should be incorporated to address the problem (RQ3).

2 Background Literature

Prior to conducting our literature review, we searched for existing review papers in this space. We identified two survey papers that categorized different nudity detection approaches. Ries et al. [11] categorized visual adult image recognition approaches to three main groups: skin-color, shape, and local feature-based approaches. They reported that color-based and shape-based approaches seemed more robust than only color-based approaches because they result in more true positive rates and less false positives rates. Shayan et al. [8] conducted a literature review on different approaches for adult image filtering techniques and pornographic image recognition. They have categorized adult image filtering techniques including keyword-based methods, IP-based blacklist methods, and visual content-based methods. They used the same three categories for pornographic image recognition approaches as Ries et al. [11]. They argued that a quantitative comparison of these approaches is not possible because of the lack of a standard dataset and definition for this field.

While these literature review papers categorized the different approaches for pornography and adult image detection, as well as some of the challenges in implementing these approaches, they did not specifically address the intersection of adolescent's online safety and machine learning for nudity and/or skin detection approaches. We argue that more human-centered approaches are needed if we are to implement such solutions in real-world situations. Therefore, we took on a more human-centric perspective for conducting our literature review on adolescent online safety and nudity detection algorithms.

Over the last decade, academic and industry researchers have tried to address the dissemination of child pornographic images indirectly through exploring various nudity and skin detection techniques. To track down illicit photos of minors, for instance, Facebook, Twitter, and Bing have worked closely with organizations such as the National Center for Missing and Exploited Children's Child Victim Identification Program [12], by designing detection and mitigation solutions. Yet, the approaches taken across academic and industry research have had their limitations. The solutions mainly focus on preventing the dissemination of child pornographic images, not proactively protecting teens before the damage has been done. As such, there is a need for more effective and teen-centered approaches to tackle the problem at the source to prevent the creation and dissemination of such imagery.

Therefore, we report key trends and potential gaps in the field of nudity and/or skin detection as it relates to the problem of adolescent sexting behaviors to bring the attention of the research community to this matter. Our research contributions include the following:

1. A synthesis of 45 peer-reviewed articles from computational literature related to nudity and/or skin detection.
2. Identification of potential gaps in the literature on computational nudity and/or skin detection techniques as it relates to adolescent sexting behaviors.
3. Recommendations for future research related to nudity and/or skin detection for the purpose of adolescent online safety.

3 Methods

We performed a comprehensive review of the nudity detection literature and identified 45 peer-reviewed articles that summarize state-of-the-art approaches. We used a grounded approach to qualitatively code the articles. We created a codebook, added codes as we read each paper, and iteratively refined the codebook throughout the coding process. Below is the description of how the systematic literature search was conducted and how the literature was synthesized.

3.1 Literature Search

We searched for articles involving adolescent sexting behaviors from mobile smartphones (both through photos and video imagery), through various libraries and databases including IEEE Xplore Digital Library, ACM Digital Library, and Springer-link to ensure comprehensive coverage of the existing literature. For this review, we focused on nudity detection using the following search terms: “teen nudity,” “adolescent nudity,” “nudity detection,” “skin detection,” “explicit content,” and “censored nudity.” We used all the mentioned resources to come up with an initial set of articles that mainly focus on the detection of nudity and explicit content. We then walked through the citations from the initial set of articles to identify additional articles that were relevant to include in our data set. We also used Google Scholar to conduct a wider search to ensure we were inclusive to multidisciplinary peer-reviewed articles.

We used the following inclusion criteria to evaluate if an article was relevant to our review: (1) The study was a peer-reviewed published work, (2) The study was published after 2008, and (3) The study must suggest a technique to detect nudity and/or skin that could potentially be used for the purpose of mitigating adolescent sexting behaviors. Articles that did not meet all three criteria were considered irrelevant and were not included in our review.

3.2 Data Analysis Approach

We identified 45 relevant articles that met the inclusion criteria mentioned above. The first author then coded them based on the dimensions found in Table 1. A grounded coding approach was used to code the articles, and the codes were iteratively updated as new codes emerged. We mapped the coded dimensions to our three high-level research questions presented in the introduction of our paper. For example, pre or post capture type (whether the algorithm ran in real-time or on a post-digitized image) and the type of media being processed (e.g., image, video, etc.) aligned with our RQ1 on the characteristics that were common to this type of research. For RQ2, we coded for the algorithmic approach used in the respective studies. For RQ3, we used a more user-centered lens to determine whether and how the research incorporated risk mitigation strategies or user studies to apply the research findings to real-world settings. The coded dimensions, their descriptions, and the corresponding codes can be found in Table 1.

Table 1. Final codebook

RQs	Dimension	Description (codes)
RQ1	Capture type	The technique detects nudity before or after capturing/digitization of image (PRE, POST)
	Media type	The technique mentioned is applicable to images, videos, both or none (IMAGE, VIDEO, GEN)
	Nudity class	The technique detects the nudity only when the subject is completely nude, private parts are naked, or it's only non-nude sexually suggestive, or other general approaches to consider nudity (COMP, PP, NN_SUGG, GEN)
	Nudity type	The technique specifically detects nudity of teens (under 18), or general nudity (TEEN, GENERIC)
	Platform type	The technique provides a solution as general software or as mobile apps and services (GEN, MOBILE)
RQ2	Detection type	The approach is for nudity detection (NUDITY) or skin detection (SKIN)
	Detection approach	The implementation makes use of machine learning techniques, computer vision or natural language processing (ML, CV, NLP)
RQ3	User study	Is the article based on a survey where a certain population is studied (YES, NO)
	Risk mitigation approach	The article suggests an implementation of blocking or the reporting of the detected content (BLOCK, REPORT)

4 Results

We present the key findings from our literature review in this section. We organize and present the results by our code dimensions. We first present the results by capture and media type, followed by nudity class and type, platform type, detection type and approach, and finally by user studies and risk mitigation.

4.1 Capture and Media Type

The code "PRE" in Table 1 for "Capture type" means that the detection of nudity happens even before the image is digitized in standard RGB matrix form. All 45 articles which we reviewed, all of them focused on post-digital capture imagery (i.e., "POST"), as opposed to pre-digital capture (i.e., "PRE"). This means that a digitized, nude image had to exist in order for the nudity detection approaches to be effective. Most of the computational academic work related to the detection of nudity has been designed to increase the accuracy and efficiency of detecting nudity [13–21], but they all operate on the already-digitally-stored instances of nudity.

Already-digitally stored images and videos can be classified very accurately through sophisticated techniques like extraction of low and high-level features [22]. This requires large processing time and space. However, when these techniques are used in real-time, it is crucial to follow time deadlines and memory constraints [23]. This indicates that detecting nudity live or in real-time applications like Skype is challenging and reliance on low-level features is more appropriate in these cases. The post-capture digital media analyzed in the reviewed articles included images (66% of articles), videos (25% of articles), text (11% of articles), and mixed media (i.e., "GEN", 23% of articles). The articles that analyzed images mostly used visual learning techniques, like feature extraction, to classify the nude or non-nude image.

Deselaers [24] proposed a method for detecting adult nudity in videos based on a bag-of-visual-features representation for frames. Kovac et al. [25] provided a method for detecting skin color based on RGB color space. Lin et al. [26] used a Support Vector Machine (SVM), which has learning skills in the image detection of human nudity. PhotoDNA [27] is one of the most popular and latest technological solutions for detecting digital nudity by analyzing digital imagery and metadata compared to a database of known images developed by Microsoft.

Though some articles [14, 16–18, 20, 24] focused on detection of explicit content in videos, few made use of the temporal correlations inside video data. Behrad et al. [15] utilized different novel features for obscene video content recognition including spatial, spatiotemporal, and motion-based features, using 3D skin volume method. Khan et al. [28] used the Viola-Jones object detection framework that works on real-time data to detect skin. Polastro et al. [20] considered the contribution of the percentage of explicit frames while classifying the video as pornographic. Hence, their techniques could also be applied to single frames that is the same as detecting nudity in images. Apart from these three papers, all who indicated their focus on video did not actually implement a technique that made use of distinctive video properties.

Text, in most cases, was analyzed on the metadata attached to the image or video under consideration. Some of the articles that we studied focused on combining already available techniques to come up with the nudity filtering system. Wijesinghe et al. [29] proposed improvement in a Parental Control and Filtering System by combining techniques like site restrictions, denial of access to web proxy servers, identification of images containing nudity, and control over uploads of images of people. We coded these articles as "GEN" (general) for media type.

4.2 Nudity Class and Type

We categorized the types of nudity detected in three classes as per their level of explicitness. We found that 40.1% of the articles proposed a method that would only detect nudity if the image was completely nude. These techniques relied more on skin detection percentage rather than region-based feature extraction, or they used some hardcoded high-level feature assumption like the use of a navel recognizing process [30]. A little over 7% of the articles presented a solution that would detect exposed private parts of one's body even if the body was mostly covered; they focused more on high and low-level feature extraction (e.g., Santos et al. [31]). Four of the articles had such a complex way of extracting features that they could almost be classified as using contextual learning for nudity. Therefore, they could detect if a naked body image was sexual or non-sexual, like in the case of a breastfeeding mother. Sevimli et al. [32] made use of 4 descriptors (feature extraction methods) to classify images in 5 classes of nudity: normal images (class 1), swimming suit images (class 2), topless images (class 3), nude images (class 4), and sexual activity images (class 5).

Most (70.5%) of the articles focused on general (primarily adult) nudity detection, as opposed to specifically detecting the nudity of a minor. In the 29.5% of articles that focused on teen nudity detection, age and nudity were separately. To determine age, very few made use of the image itself; rather they used additional information like text from metadata or information entered directly by the uploader of the image. Articles that propose a solution to general nudity detection in terms of age mostly make use of two very general steps: skin detection and pornography detection [33, 34]. For detecting teen nudity, the third step, 'age detection' is added. Polastro et al. [35] makes it clear that there are some intrinsic characteristics that need to be considered to distinguish child nudity from adult nudity. For example, most of the child sexual content does not contain explicit sex scenes and the motion (in a video) can be distinguished from an adult porn scene. Also, the sound can be distinguished since it is different or sometimes absent in the case of child porn video. Figure 1 depicts the number of articles for methods specific for teens and general methods and their nudity class.

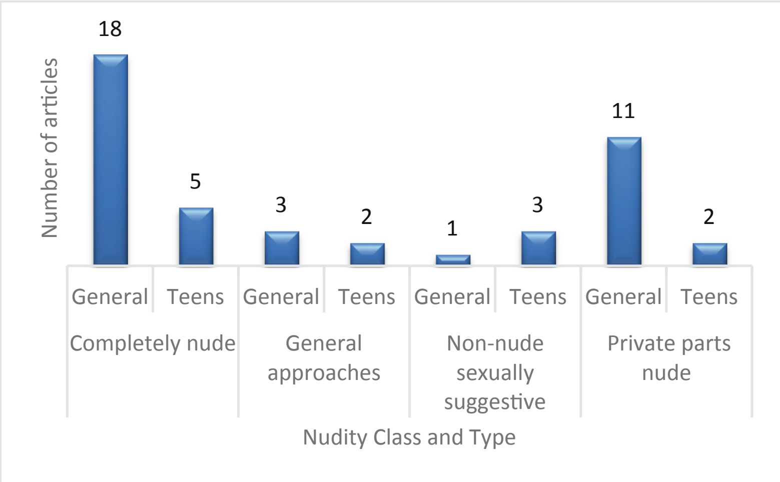

Fig. 1. Number of articles by nudity class and type

4.3 Platform Type

Even though mobile social networking and dating apps are a major platform for sexting [36], only 5% of the articles focused on detecting nude content specifically via mobile devices. While the software-based detection techniques that most of the researchers presented were general enough to be used via mobile platforms, none were tested to see if the mobile processor had the adequate processing power for the detection task. Two of the articles present a third-party mobile-application to detect explicit content. Lin et al. [37] presented an app that detects skin and classifies it, and for further detection of nudity, the app would conduct a user poll.

Amato et al. [38] presents an application that is more of a background service, an interceptor that detects nudity or other explicit content being received through MMS and Bluetooth messaging on mobile devices based on the Symbian™ operating systems. Once intercepted, the images are analyzed by the component of the system that automatically classifies images with explicit sexual content. Apart from these two articles, all others provide a general solution that can be implemented at a different level of abstraction of the operating systems, depending on the authority. For instance, solutions could be implemented by the Internet Service Provider at the network level to provide information to law enforcement agencies or by an application that is controlled by parents of the device.

4.4 Detection Type and Approach

In order to detect nude scenes, detection of different kinds of contextual and visual features in an image are necessary. In reviewing the literature, skin was one of the important features for CV nudity detection techniques [39, 40]. Islam et al. [41] stated that "Nudity and pornography have a direct link with human skin. In fact, no pornography can exist without exposure of human skin. Apart from pornography, a wide range of image processing applications exist, where skin detection is playing a crucial role. Using color as a detection cue has long being recognized as a robust feature and has become a popular choice in human skin detection techniques. Human skin has a characteristic color which is easily distinguishable from the colors of other objects."

Nine of the papers [23, 28, 32, 34, 41–46] implemented skin detection as a part of the procedure to detect nudity. Islam et al. [41] used a Wavelet transform that involves recursive filtering and sub-sampling. It has discriminating ability in texture analysis that facilitates capturing subtle differences between child and adult skin texture. Bhoyar [23] proposed a three-layer feedforward neural network used for skin color classification with three neurons in the input layer, five neurons in the hidden layer and two neurons in the output layer. The two neurons in the output layer represent skin class and non-skin class. Povar et al. [42] and Kelly et al. [43] used clustering in color space (s) to filter skin tone. Sevimli et al. [32] used a method based on inferring pixels on statistical skin and non-skin models which are represented and trained with Gaussian Mixture Models. Vijayendar et al. [47] proposed a method to filter adult images in websites; they use MFC(Most Frequent Color) where face detection is not possible. Selamat et al. [48] used modified fuzzy rules to improve skin detection. Choudhury et al. [49] proposed a skin tone detection filter that can identify images with a large skin color count that are pornographic in nature. Dewantono et al. [50] proposed nudity detection and localization in images and videos using a skin filtering method based on a Bayes rule, a novel histogram back projection of skin samples, and a SVM. Siqueira et al. [34] constructed color histograms through both the skin and non-skin groups of RGB images; they applied a certain threshold on the histograms to classify pixels into groups. Khan et al. [28] used adaptive skin color modeling where pixels that are most likely non-skin are discarded from a detected region of pixels and the region is then extracted for further processing. Liu et al. [44] proposes to detect pornographic images in a two-stage scheme; the first step employs a Content-Based Image Retrieval technique (CBIR) to determine whether the image has a human in it. The second step is a skin color model established to analyze the skin-like pixels and identify the presence of pornographic content.

Soysal et al. [51] proposed a concept detection system (one of the concepts is nudity detection) using generalized visual and audio concept detection modules. Adnan et al. [16] analyzed low-level features for their suitability in pornography detection; they found that in order for pornography detection systems to be accurate, not only do low-level features need to be considered, but high-level features should also be incorporated. Esposito et al. [52] proposed a nudity detection classifier based on both body geometric properties and global features. Eleuterio et al. [53] improved child pornography video detection algorithms by proposing an adaptive sampling approach.

Westlake et al. [54] tackled the dissemination of online child sexual exploitation (CE) using a different approach; they investigated the communities that are created around public websites involved in the distribution of child sexual exploitation material. One criterion of identifying a website as CE-related was that it contained one of a set of known images from Royal Canadian Mounted Police (RCMP) database using a hash value. Girgis et al. [55] proposed a pornography detection and filtering system for images in web pages using skin recognition. In addition, CBIR methods can be used for nudity detection filters [56].

From a computational perspective, most articles (86.4%) employed methods from the field of Computer Vision (CV) [50, 52, 57–61]. Only 9% of the articles proposed Natural Language Processing (NLP) techniques. A few papers presented pure CV solutions without any learning aspect; for example, Ivan et al. [62] used the RSOR algorithm that performed recognition and selection of the largest region in a segmented image. Some articles (e.g., Polastro et al. [63]) used mixed techniques from ML, CV, and NLP to come up with a comprehensive technique to detect nudity. Vanhove et al. [14] presented a solution that used Picture Analysis, Text Analysis, and Audio/Video analysis for social network monitoring. More detailed categorizations of different pornography and adult image detection methods could be found in Ries et al. and Shayan et al. articles [8, 11].

4.5 User Studies and Risk Mitigation Strategies

A common theme among the nudity detection approaches was that they came from a purely computational perspective and failed to incorporate any aspects of user-centered design or needs analysis. None of the articles included formative or summative user evaluations of the solutions developed. For instance, an online article [64] talks about a parental control software and briefly mentions mitigation approaches that require a user study, which can consider different mitigation approaches to the problem depending on the relation or trust between parents and children. In terms of risk mitigation strategies after nudity detection has occurred, the article set was also lacking. About 22.7% of the articles suggested naïve approaches, such as blocking the explicit content or reporting it. In another 27%, the articles suggested that their solutions could be used by law enforcement agencies, security companies, or parents. None of the articles suggested any kind of design that would directly engage teen users in a way to address the root of the risky behavior. In summary, the majority of the articles focused on risk detection over risk mitigation.

Table 2 summarizes all articles with the research issues that they tried to address. Most of the articles proposed nudity and/or skin detection systems without embedding them in any particular context or application. Though some of the papers proposed parental control apps and/or filtering systems, other articles proposed cataloging tools and/or retrieval systems. These systems could be used by different agencies, such as police investigation departments. There were no cases were the articles proposed that the technologies were suited for use in applications targeted for teen users.

Table 2. Source articles and their intended purpose

Topic/purpose	Sources
Skin detection	Islam et al., 2011 [41]; Bhoyar et al., 2010 [23]; Siqueira et al., 2013 [34]; Selamat et al., 2009 [48]; Santos et al., 2015 [46], Kelly et al., 2008 [43]
Nudity detection	Flores et al., 2011 [62]; Santos et al., 2012 [31], Polastro et al., 2012 [35]; Soysal et al., 2013 [51]; Behrad et al., 2012 [15]; Adnan et al., 2016 [16]; Esposito et al., 2013 [52]; Sevimli et al., 2010 [32]; Uke et al., 2012 [17]; Silva et al., 2014 [18]; Lin et al., 2012 [37]; Ras et al., 2016 [61]; Platzer et al., 2014 [33]; Eleuterio et al., 2010 [63]; Polastro et al., 2012 [20]; Dewantono et al., 2014 [50]; Deselaers et al., 2008 [24]; Wang et al., 2009a [65]; Wang et al., 2009b [30]; Lopes et al.,2009 [21]; Ap-apid, 2009 [45]; Liu et al., 2009 [44]; Eleuterio et al., 2012 [53]; Lopes et al., 2009 [66]; Steel et al., 2012 [60]; Polastro et al., 2010 [63]
Parental control and/or filtering system	Wijesinghe et al., 2012 [29]; Ahuja et al., 2015 [29]; Lienhart et al., 2009 [57]; Vijayendar et al., 2009 [47]; Amato et al., 2009 [38]; Khan et al., 2008 [28]; Girgis et al., 2010 [55]; Vanhove et al., 2013 [14]; Choudhury et al., 2008 [49]
Cataloguing tool and/or retrieval system	Povar et al., 2011 [42]; Grega et al., 2011 [67]; Patil et al., 2013 [22]; Sidhu et al., 2015 [56]

5 Discussion

Our review indicated potential gaps in the literature that can inform new research directions. In this section, we summarize our major findings and identify the gaps and opportunities for future research on nudity and/or skin detection for the purpose of mitigating risks associated with adolescent sexting behaviors.

5.1 Summary of Findings

Most research related to the detection of nudity and/or skin has been studied at the software-level on already digitized images. A problem with this approach is that it implies that the naked image of a minor must already be digitized in order for the algorithm to work. This raises a crucial need for a more effective approach that addresses the problem at the source to prevent the creation and dissemination of such imagery in the first place. Skin detection at the pre-digitization level might ensure privacy and, in the future, can be combined with detection of other spatial and/or temporal features in nude scenes to prevent the digitization of such images at all.

We also found limited research related to the detection of nudity and/or skin detection within the context of mobile devices. Since most teens use mobile devices as their primary means for going online [1], more research in this area is warranted. Researchers should focus on addressing the limitations of processing power and memory to apply their general solutions to the nudity and/or skin detection problem within mobile platforms.

Another limitation we found within the articles was that very few researchers contextualized their algorithmic solutions to children and teens. Many of the studies did not use teen data sets to train the nudity and/or skin detection algorithms, nor did they validate ground truth based on adolescents. Additionally, many of the algorithms themselves are often generic and not optimized for adolescent data. The algorithms are typically developed outside of the context of how they would need to be used for adolescent sexting risk detection. Therefore, researchers should consider conducting studies with teen datasets. These studies can help determine the level of nudity that constitutes risks for teens to determine thresholds that could establish a ground truth in nudity and/or skin detection for adolescent online safety.

Finally, we found that articles focus more on risk detection, rather than risk mitigation. Future research should shift the focus to more proactive solutions by incorporating aspects of user-centered design or developing formative and summative user evaluations of the solutions. This would allow for the design and development of interventions that would directly engage teen users in a way to actively manage their online behavior and address the root of the risky behavior.

We further discuss these gaps in the following sections.

5.2 The Need to Detect Nudity and/or Skin Before Digital Capture

To our knowledge, there is no peer-reviewed literature that deals with the issue of handling teen sexting behaviors before they have already been digitally captured via a nude image or video. Skin detection at the pre-digitization level has the advantage of privacy, and in the future, it can be combined with detection of other spatial and/or temporal features in nude scenes to prevent teen sexting. The literature presents techniques that detect nudity and/or skin in digitized images or frames of videos. Unfortunately, by this point, the damage has already potentially been done. The image or video could have already reached a multitude of individuals and platforms. Also, there is no method of knowing whether the image or video was saved, allowing for the possibility of it to re-surface in the future. Therefore, a mechanism to detect nudity and/or skin at the pre-digitization level can be more useful. Two major potential benefits of pre-digitization detection include the following:

(1) **Security:** The British Broadcasting Corporation (BBC) reported that once an Apple iPhone, iPod Touch, or iPad owner grants permission for an application to access location information from their device, the application can potentially copy their photo library [33]. Also, in several high-profile examples, celebrity photos

have been leaked onto the Internet after their phones were hacked. By detecting nudity and/or skin before an image or video is captured, the probabilities of these types of risks taking place can be reduced.

(2) **Real-time detection:** At the heart of almost all techniques provided in the literature is a basic step of feature extraction. So far, there has been more emphasis on accuracy of classification of nude images which improves by extracting both low and high-level features. This requires considerable processing time and space, whereas, dealing with only low-level features will require less processing time and that would help to detect nudity and/or skin in real time with a compromise on accuracy. However, nudity detection in live or real-time video applications, like Skype, could be possible.

For example, Tariq et al. [68] developed a low-powered sensor as a proof-of-concept for detecting skin before an image is captured. They were able to detect skin dominance with 83.7% accuracy. Future research should build upon this work and investigate how to develop even more effective methods of detecting nudity and/or skin at the pre-digitization level.

5.3 The Need to Develop Risk Detection Solutions for Mobile Platforms

Most articles we reviewed made use of machine learning, on the features from images and videos. Classifying them accurately requires adequate memory and processing time, which likely requires high-processing power on a centralized server. This implies that the image would first need to be sent to the centralized server before the algorithms could be applied to detect skin or nudity.

Research has shown that teens are increasingly becoming more mobile with their online access [1]. Yet, the articles we found do not specify the application of these general methods to mobile platforms, because of the limited memory and processing power of the mobile device. Amato et al. [38] mentions that "classifiers can recognize and discriminate between harmless and offensive multimedia contents (in a mobile device). However, the complexity of such systems discourages from implementing and running them on small mobile devices." On the other hand, every device has the connectivity capabilities necessary for sending and receiving a relatively rich amount of information which gives software developers the liberty to do extensive calculations at the server side. However, interruption in connectivity could be a problem especially when dealing with real time detection. For general risk detection solutions to be effective, they must be able to be applied within the appropriate context (in this case mobile devices). Therefore, researchers should consider investigating approaches for incorporating risk detection into mobile platforms.

5.4 The Need to Contextualize Risk Detection Algorithms to Adolescent Sexting Behaviors

Most of the articles that we studied propose nudity and/or skin detection methods irrespective of the subject's age. General solutions typically make use of textual data attached with the image to identify the age of the subject. However, visual learning techniques could be used to detect visual features that could identify the age range of the subjects, whether they are a child or a teen [35]. Further, contextual learning could be used to detect images that are sexually suggestive, as almost all the detection techniques mentioned above simply detect complete nudity or the private areas of a human body. These proposed methods could help prevent false positives [63], for example, correctly identifying a semi-nude sexual image of a child versus incorrectly identifying a child playing at the beach in their bathing suit. Both examples may have the same amount of nudity, but the human context would indicate whether it is an appropriate image or not.

Additionally, as far as adolescent online risk is concerned, there has not been a study to understand what level of nudity actually constitutes risks for teens. For instance, a bathroom selfie in a towel may exhibit less skin, but the suggestiveness of the context would make it riskier. As such, detecting nudity in binary terms may not adequately serve to mitigate risks. More studies should focus on understanding the thresholds used to determine ground truth in nudity detection. As discussed earlier, there was also no article, to our knowledge, that suggested a design that would directly engage teen users in a way to address the root of the risky behavior. In short, most articles focused on risk detection over risk mitigation, failing to incorporate any aspects of user-centered design or a formative/summative user evaluation of the solutions developed [69]. The detriment of this gap is made evident in existing research on the limitations of parental control applications [70, 71], which implement many of the risk detection methods discussed in this paper. Parent-focused interventions that increase parental control through restriction and monitoring may not be as effective as teen-centric solutions that empower teens to make better online decision [70–72]. Researchers should focus their efforts on designing more teen-centric risk detection algorithms that take into consideration the unique characteristics of teens.

5.5 The Need for Human-Centered Machine Learning Research

Almost all the reviewed articles were algorithms and/or systems developed by engineers and computer scientists, who did not take into consideration users' interpretations of results which might be different from their personal interpretations of results. Most of the reviewed papers used metrics such as accuracy, recall, precision, and F1 score to evaluate the performance of algorithms, but they did not incorporate any human-centered evaluation based on the perceptions of actual users. The algorithmic systems for nudity and/or skin detection need to benefit human-centered designs, approaches, and evaluations, to devise and improve the systems for adolescents to meet their specific characteristics. To meet this objective, researchers should involve adolescents and their parents to help them design and evaluate their systems [71, 73–75].

Researchers should take into consideration how machine learning algorithms may unintentionally be influenced by biases [76]. Human-centered machine learning takes this into account by considering the goals and capabilities of humans and having them help with the design and evaluation of these algorithms and systems [77]. Baumer [78] proposed a human-centered algorithm design which requires the design process for algorithmically based systems to incorporate human and social interpretations. He provided theoretical, participatory, and speculative strategies for these types of designs.

As machine learning systems are being used in more real-word systems, it is important that they are useful to people. That's where human-centered approaches could help improve human experiences in development of new machine learning technologies.

5.6 Limitations and Future Research

In the future, researchers should take the necessary steps towards a more cohesive solution for providing risk mitigation after detection. If it were possible to detect risky online behaviors (e.g., a teen taking a nude photo or streaming video while unclothed) using a teen's internet-connected device (e.g., mobile smartphone, tablet, or laptop), then we would be able to mitigate these risks in more meaningful ways. Unfortunately, nudity detection poses additional risks to teens, as a high-fidelity digitized nude image of a minor (possibly transmitted to a server for additional processing) already negates our goal of preserving the privacy of minors. Therefore, an integral part of this long-term goal of detecting nudity prior to digital capture is a sensor that integrates directly with a mobile application to decouple skin detection (performed by the sensor) from risk mitigation strategies (managed by the application layer), so that parents can customize how to handle problematic behavior based on the age and unique needs of their teen. Whether risk mitigation comes in the form of blocking digital transmission, notifying parents, or nudging teens to make better choices, detection without mitigation cannot serve to effectively protect teens.

6 Conclusion

In this paper, we synthesized past research on nudity and/or skin detection for the purpose of combatting adolescent sexting behaviors. Our review uncovers potential gaps in the literature that can inform new research directions. This work contributed the first literature review of nudity and/or skin detection with the focus of solutions customized for adolescents. The key points that future researchers should consider include the following: (1) To design nudity detection solutions at the pre-digitization levels to ensure privacy. (2) To provide solutions that are applicable on mobile platforms. (3) To conduct studies to understand specific characteristics of adolescent sexting. By designing solutions for nudity detection focusing on adolescents as end users, researchers will be able to provide more effective solutions for combating adolescent sexting behaviors.

Acknowledgements. This research was supported by the William T. Grant Foundation under grant number #187941 and the U.S. National Science Foundation under grant number #IIP-1827700. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the research sponsors.

References

1. Teens, Social Media & Technology 2018—Pew Research Center (2018). http://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/
2. Lievens, E.: Bullying and sexting in social networks: protecting minors from criminal acts or empowering minors to cope with risky behaviour? Int. J. Law Crime Justice **42**, 251–270 (2014). https://doi.org/10.1016/j.ijlcj.2014.02.001
3. Citizen's Guide To U.S. Federal Law On Child Pornography. https://www.justice.gov/criminal-ceos/citizens-guide-us-federal-law-child-pornography
4. Drouin, M., Ross, J., Tobin, E.: Sexting: a new, digital vehicle for intimate partner aggression? Comput. Hum. Behav. **50**, 197–204 (2015). https://doi.org/10.1016/j.chb.2015.04.001
5. Teens and Sexting—Pew Research Center (2009). http://www.pewinternet.org/2009/12/15/teens-and-sexting/
6. Ngejane, C.H., Mabuza-Hocquet, G., Eloff, J.H.P., Lefophane, S.: Mitigating online sexual grooming cybercrime on social media using machine learning: a desktop survey. In: 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), pp. 1–6 (2018)
7. Snapchat sexting and the predators of Kik: The apps your children need to stay away from—Daily Telegraph. https://www.dailytelegraph.com.au/news/snapchat-sexting-and-the-predators-of-kik-the-apps-your-children-need-to-stay-away-from/news-story/ad63dc7962ff4098b0785d2671f5c89c
8. Shayan, J., Abdullah, S.M., Karamizadeh, S.: An overview of objectionable image detection. In: 2015 International Symposium on Technology Management and Emerging Technologies (ISTMET), pp. 396–400 (2015)
9. Lee, S.-J., Chae, Y.-G.: Balancing participation and risks in children's Internet use: the role of internet literacy and parental mediation. Cyberpsychol. Behav. Soc. Netw. **15**, 257–262 (2012). https://doi.org/10.1089/cyber.2011.0552
10. Byrne, S., Katz, S.J., Lee, T., Linz, D., McIlrath, M.: Peers, predators, and porn: predicting parental underestimation of children's risky online experiences. J. Comput.-Mediat. Commun. **19**, 215–231 (2014). https://doi.org/10.1111/jcc4.12040
11. Ries, C.X., Lienhart, R.: A survey on visual adult image recognition. Multimed Tools Appl. **69**, 661–688 (2014). https://doi.org/10.1007/s11042-012-1132-y
12. PhotoDNA Lets Google, FB and Others Hunt Down Child Pornography Without Looking at Your Photos. https://petapixel.com/2014/08/08/photodna-lets-google-facebook-others-hunt-down-child-pornography-without-looking-at-your-photos/
13. Raverkar, S.D., Nagori, M.: Classification of YouTube metadata using shark algorithm. Int. J. Comput. Appl. **132**, 18–21 (2015)
14. Vanhove, T., Leroux, P., Wauters, T., Turck, F.D.: Towards the design of a platform for abuse detection in OSNs using Multimedial data analysis, p. 4 (2013)
15. Behrad, A., Salehpour, M., Ghaderian, M., Saiedi, M., Barati, M.N.: Content-based obscene video recognition by combining 3D spatiotemporal and motion-based features. J Image Video Proc. **2012**, 23 (2012). https://doi.org/10.1186/1687-5281-2012-23
16. Adnan, A., Nawaz, M.: RGB and hue color in pornography detection. Information Technology: New Generations. AISC, vol. 448, pp. 1041–1050. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32467-8_90
17. Uke, N.J., Thool, D.R.C.: Detecting pornography on web to prevent child abuse – a computer vision approach. Int. J. Sci. Eng. Res. **3**, 3 (2012)

18. Silva, P.M., Polastro, M.D.: An overview of NuDetective forensic tool and its usage to combat child pornography in Brazil. Presented at the XLIII Jornadas Argentinas de Informática e Investigación Operativa (43JAIIO)-VI Workshop de Seguridad Informática (Buenos Aires, 2014) (2014)
19. Eleuterio, P., Polastro, M.: Identification of high-resolution images of child and adolescent pornography at crime scenes. Int. J. Forensic Comput. Sci., 49–59 (2010). https://doi.org/10.5769/j201001006
20. Polastro, M., Eleuterio, P.: Quick identification of child pornography in digital videos. Int. J. Forensice Comput. Sci. **7**, 21–32 (2012). https://doi.org/10.5769/J201202002
21. Lopes, A.P., de Avila, S.E., Peixoto, A.N., Oliveira, R.S., Coelho, M.D., Araújo, A.D.: Nude detection in video using bag-of-visual-features. In: 2009 XXII Brazilian Symposium on Computer Graphics and Image Processing, pp. 224–231 (2009)
22. Patil, S.T., Chavan, N.A.: Image cataloguing tool using descriptor for forensic application. Int. J. Comput. Appl. **975**, 8887 (2013)
23. Bhoyar, K.K., Kakde, O.G.: Skin color detection model using neural networks and its performance evaluation. J. Comput. Sci. **6**, 963–968 (2010)
24. Deselaers, T., Pimenidis, L., Ney, H.: Bag-of-visual-words models for adult image classification and filtering. In: 2008 19th International Conference on Pattern Recognition, pp. 1–4 (2008)
25. Kovac, J., Peer, P., Solina, F.: Human skin color clustering for face detection. In: The IEEE Region 8 EUROCON 2003. Computer as a Tool, pp. 144–148, vol. 2 (2003)
26. Lin, Y.-C., Tseng, H.-W., Fuh, C.-S.: Pornography detection using support vector machine, p. 8 (2003)
27. PhotoDNA—Microsoft. https://www.microsoft.com/en-us/photodna
28. Khan, R., Stöttinger, J., Kampel, M.: An adaptive multiple model approach for fast content-based skin detection in on-line videos. In: Proceedings of the 1st ACM Workshop on Analysis and Retrieval of Events/Actions and Workflows in Video Streams, pp. 89–96. ACM, New York (2008)
29. Wijesinghe, S., Wijewardana, V.O., Karunarathna, Y.D.D., Ridmal, H.S.: Parental Control and Filtering System (2013)
30. Wang, X., Hu, C., Yao, S.: An adult image recognizing algorithm based on naked body detection. In: 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, pp. 197–200 (2009)
31. Santos, C., dos Santos, E.M., Souto, E.: Nudity detection based on image zoning. In: 2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), pp. 1098–1103 (2012)
32. Sevimli, H., et al.: Adult image content classification using global features and skin region detection. In: Gelenbe, E., Lent, R., Sakellari, G., Sacan, A., Toroslu, H., Yazici, A. (eds.) Computer and Information Sciences, pp. 253–258. Springer, Netherlands (2010). https://doi.org/10.1007/978-90-481-9794-1_49
33. Platzer, C., Stuetz, M., Lindorfer, M.: Skin sheriff: a machine learning solution for detecting explicit images. In: Proceedings of the 2nd International Workshop on Security and forensics in Communication Systems - SFCS 2014, pp. 45–56. ACM Press, Kyoto (2014)
34. de Siqueira, F.R., Schwartz, W.R., Pedrini, H.: Adaptive Detection of Human Skin in Color Images. 5
35. de Castro Polastro, M., da Silva Eleuterio, P.M.: A statistical approach for identifying videos of child pornography at crime scenes. In: 2012 Seventh International Conference on Availability, Reliability and Security, pp. 604–612 (2012)
36. Apps to improve your sext life - NY Daily News. http://www.nydailynews.com/news/national/9-apps-improve-sext-life-article-1.2465165

37. Lin, Y., Wu, Y.: Machine learning application on detecting nudity in images. Presented at the (2012)
38. Amato, G., Bolettieri, P., Costa, G., la Torre, F., Martinelli, F.: Detection of images with adult content for parental control on mobile devices? In: Proceedings of the 6th International Conference on Mobile Technology, Application & Systems, pp. 35:1–35:5. ACM, New York (2009)
39. Smola - Introduction to Machine Learning.pdf. https://alex.smola.org/teaching/cmu2013-10-701/slides/3_Instance_Based.pdf
40. A Hierarchical Method for Nude Image Filtering– 《Journal of Computer Aided Design & Computer Graphics》 2002年05期. http://en.cnki.com.cn/Article_en/CJFDTOTAL-JSJF200205004.htm
41. Islam, M., Watters, P.A., Yearwood, J.: Real-time detection of children's skin on social networking sites using Markov random field modelling. Inf. Secur. Tech. Rep. **16**, 51–58 (2011). https://doi.org/10.1016/j.istr.2011.09.004
42. Povar, D., Vidyadharan, D.S., Thomas, K.L.: Digital image evidence detection based on skin tone filtering technique. In: Abraham, A., Lloret Mauri, J., Buford, J.F., Suzuki, J., Thampi, S.M. (eds.) ACC 2011. CCIS, vol. 190, pp. 544–551. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22709-7_53
43. Kelly, W., Donnellan, A., Molloy, D.: Screening for objectionable images: a review of skin detection techniques. In: 2008 International Machine Vision and Image Processing Conference, pp. 151–158 (2008)
44. Liu, B., Su, J., Lu, Z., Li, Z.: Pornographic images detection based on CBIR and skin analysis. In: 2008 Fourth International Conference on Semantics, Knowledge and Grid, pp. 487–488 (2008)
45. Ap-apid, R.: Image-based pornography detection. Presented at the January 1 (2008)
46. Santos, A., Pedrini, H.: Human skin segmentation improved by saliency detection. In: Azzopardi, G., Petkov, N. (eds.) CAIP 2015. LNCS, vol. 9257, pp. 146–157. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23117-4_13
47. Vijayendar G.: Integrated approach to block adult images in websites. In: 2009 International Conference on Computer Technology and Development, pp. 421–425 (2009)
48. Selamat, A., Maarof, M.A., Chin, T.Y.: Fuzzy mamdani inference system skin detection. In: 2009 Ninth International Conference on Hybrid Intelligent Systems, pp. 57–62 (2009)
49. Choudhury, A., Rogers, M., Gillam, B., Watson, K.: A novel skin tone detection algorithm for contraband image analysis. In: 2008 Third International Workshop on Systematic Approaches to Digital Forensic Engineering, pp. 3–9 (2008)
50. Dewantono, S., Supriana, I.: Development of a real-time nudity censorship system on images. In: 2014 2nd International Conference on Information and Communication Technology (ICoICT), pp. 30–35 (2014)
51. Soysal, M., et al.: Multimodal concept detection in broadcast media: KavTan. Multimedia Tools Appl. **72**, 2787–2832 (2014). https://doi.org/10.1007/s11042-013-1564-z
52. Esposito, L.G., Sansone, C.: A multiple classifier approach for detecting naked human bodies in images. In: Petrosino, A. (ed.) ICIAP 2013. LNCS, vol. 8157, pp. 389–398. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41184-7_40
53. Eleuterio, P., Polastro, M.: An adaptive sampling strategy for automatic detection of child pornographic videos. In: Proceedings of the Seventh International Conference on Forensic Computer Science, pp. 12–19. Abeat (2012)
54. Westlake, B.G., Bouchard, M.: Liking and hyperlinking: community detection in online child sexual exploitation networks. Soc. Sci. Res. **59**, 23–36 (2016). https://doi.org/10.1016/j.ssresearch.2016.04.010

55. Girgis, M.R., Mahmoud, T.M., Abd-El-Hafeez, T.: A new effective system for filtering pornography images from web pages and PDF files. IJWA **2**, 13 (2010)
56. Sidhu, S., Saxena, J.: Content Based Image Retrieval A Review, 5 (2015)
57. Lienhart, R., Hauke, R.: Filtering adult image content with topic models. In: Proceedings of the 2009 IEEE International Conference on Multimedia and Expo, pp. 1472–1475. IEEE Press, Piscataway (2009)
58. Nude image detection based on SVM. https://www.researchgate.net/publication/224586727_Nude_image_detection_based_on_SVM
59. Ahuja, C., Baghel, A.S., Singh, G.: Detection of nude images on large scale using Hadoop. In: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 849–853 (2015)
60. Steel, C.M.S.: The mask-SIFT cascading classifier for pornography detection. In: World Congress on Internet Security (WorldCIS-2012), pp. 139–142 (2012)
61. Ras, L.T., Barfeh, D.P.Y.: Nudity prohibition system. LPU-Laguna J. Eng. Comput. Stud. **3** (2016)
62. Flores, P.I.T., Guillén, L.E.C., Prieto, O.A.N.: Approach of RSOR algorithm using HSV color model for nude detection in digital images. Comput. Inf. Sci. **4**, 29. https://doi.org/10.5539/cis.v4n4p29
63. de_Castro Polastro, M., da Silva Eleuterio, P.M.: NuDetective: a forensic tool to help combat child pornography through automatic nudity detection. In: 2010 Workshops on Database and Expert Systems Applications, pp. 349–353 (2010)
64. Discipline for Defiant Teens and Preteens. http://www.myoutofcontrolteen.com/
65. Xin-Lu, W., Xiao-juan, L., Xiao-bo, L.: Nude image detection based on SVM. In: 2009 International Conference on Computational Intelligence and Natural Computing, pp. 178–181 (2009)
66. Lopes, A., Avila, S., Peixoto, A., Oliveira, R., Araujo, A.: A Bag-of-features Approach based on Hue-SIFT Descriptor for Nude Detection. 5
67. Grega, M., Bryk, D., Napora, M.: INACT—INDECT advanced image cataloguing tool. Multimed Tools Appl. **68**, 95–110 (2014). https://doi.org/10.1007/s11042-012-1164-3
68. Tariq, M.U., Ghosh, A.K., Badillo-Urquiola, K., Jha, A., Koppal, S., Wisniewski, P.J.: Designing light filters to detect skin using a low-powered sensor. In: SoutheastCon 2018, pp. 1–8 (2018)
69. Pinter, A., Wisniewski, P., Xu, H., Rosson, M.B., Carroll, J.M.: Adolescent online safety: moving beyond formative evaluations to designing solutions for the future. In: the 16th Interactive Design and Children Conference, Stanford, CA (2017)
70. Ghosh, A.K., Badillo-Urquiola, K., Rosson, M.B., Xu, H., Carroll, J.M., Wisniewski, P.J.: A matter of control or safety?: Examining Parental use of technical monitoring apps on teens' mobile devices. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 194:1–194:14. ACM, New York (2018)
71. Ghosh, A., Badillo-Urquiola, K., Guha, S., LaViola, J., Wisniewski, P.: Safety vs. surveillance: what children have to say about mobile apps for parental control. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (2018)
72. Wisniewski, P., et al.: Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 4029–4038. ACM, New York (2015)
73. Badillo-Urquiola, K., Smriti, D., McNally, B., Bonsignore, E., Golub, E., Wisniewski, P.: Co-designing with children to address "stranger danger" on Musical.ly. Presented at the SOUPS, The Fourteenth Symposium on Usable Privacy and Security (2018)

74. Wisniewski, P., Xu, H., Rosson, M.B., Perkins, D.F., Carroll, J.M.: Dear diary: teens reflect on their weekly online risk experiences. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 3919–3930. ACM (2016)
75. Wisniewski, P., Xu, H., Rosson, M.B., Carroll, J.M.: Parents just don't understand: why teens don't talk to parents about their online risk experiences. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 523–540. ACM (2017)
76. Barocas, S., Selbst, A.D.: Big data's disparate impact. Calif. L. Rev. **104**, 671 (2016)
77. Fiebrink, R., Gillies, M.: Introduction to the special issue on human-centered machine learning. ACM Trans. Interact. Intell. Syst. **8**, 7:1–7:7 (2018). https://doi.org/10.1145/3205942
78. Baumer, E.P.: Toward human-centered algorithm design. Big Data Soc. **4**, (2017). https://doi.org/10.1177/2053951717718854