The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition

OS KEYES, University of Washington, USA

Automatic Gender Recognition (AGR) is a subfield of facial recognition that aims to algorithmically identify the gender of individuals from photographs or videos. In wider society the technology has proposed applications in physical access control, data analytics and advertising. Within academia, it is already used in the field of Human-Computer Interaction (HCI) to analyse social media usage. Given the long-running critiques of HCI for failing to consider and include transgender (trans) perspectives in research, and the potential implications of AGR for trans people if deployed, I sought to understand how AGR and HCI understand the term "gender", and how HCI describes and deploys gender recognition technology. Using a content analysis of papers from both fields, I show that AGR consistently operationalises gender in a trans-exclusive way, and consequently carries disproportionate risk for trans people subject to it. In addition, I use the dearth of discussion of this in HCI papers that apply AGR to discuss how HCI operationalises gender, and the implications that this has for the field's research. I conclude with recommendations for alternatives to AGR, and some ideas for how HCI can work towards a more effective and trans-inclusive treatment of gender.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; • **Social and professional topics** → **Socio-technical systems**; **Gender**; *Surveillance*; • **Information systems** → *Computational advertising*; • **Computing methodologies** → *Computer vision*; • **Security and privacy** → Social aspects of security and privacy;

Additional Key Words and Phrases: automatic gender recognition; gender; machine learning; transgender

ACM Reference Format:
Os Keyes. 2018. The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 88 (November 2018), 22 pages. https://doi.org/10.1145/3274357

1 INTRODUCTION

Researchers within Human-Computer Interaction have long studied the way that design processes dominated by men produce gendered, material differences in the usability of the resulting artefacts for women. But this frame of study is limited: too often, HCI research has implicitly or explicitly treated "gender" as a binary, immutable and physiologically-discernible concept. Such a model fundamentally erases transgender people, excluding their concerns, needs and existences from both design and research. The consequence has been a tremendous underrepresentation of transgender people in the literature, recreating discrimination found in the wider world (which, in the West, traditionally relies on a similar concept of gender).

In this paper, I explore how gender is conceived of and reproduced (or "operationalised") in Automatic Gender Recognition (AGR), a sub-field of facial recognition that seeks to algorithmically identify gender from photographs. AGR has a variety of societal applications, including physical

Author's address: Os Keyes, University of Washington, 428 Sieg Hall, Box 352315, Seattle, WA, 98195-2315, USA, okeyes@uw.edu.

2573-0142/2018/11-ART88 $15.00
https://doi.org/10.1145/3274357

access control and analytics, and has already been used in HCI research around social media. My goals are twofold. First, I wish to examine AGR's treatment of gender, and what the consequences of this are likely to be for transgender people as the technology is deployed. Second, I wish to analyse how HCI frames this technology as it accepts it, and use this as a lens to understand whether progress has been made in the trans-inclusivity of the field. This work thus examines the assumptions and ethical underpinnings of part of the field's research tools, and investigates the wider societal implications of this kind of algorithm.

I undertake this exploration by conducting a content analysis of prominent AGR literature, and of HCI papers that rely on the technology, asking two questions:

(1) How does Automated Gender Recognition research operationalise gender, and what are the possible consequences of this should it be widely deployed?
(2) How does HCI research interacting with AGR operationalise gender and contextualise any gendered assumptions of AGR software?

My findings show that AGR research fundamentally ignores the existence of transgender people, with dangerous results. An analysis of the HCI papers that rely on the technology reveals a similar operationalisation, with similar erasure, and no discussion of the issues with AGR. Taken in conjunction with other work in HCI, this suggests that the way in which gender is commonly operationalised in the field is one cause of HCI's erasure of trans users' needs and experiences.

I conclude by discussing alternative approaches and methodologies for designers or researchers considering AGR, and how HCI might as a community go about developing and using a more nuanced operationalisation of gender. I believe that doing so is a first step towards a more trans-inclusive approach to research, and consequently more equitable outcomes from the work HCI does.

2 BACKGROUND

2.1 Gender in society

Traditionally, Western culture has alternately conflated and drawn arbitrary distinctions between two constructs: sex, a person's biological category (male or female) based on anatomy, chromosomes and hormones, and gender, a person's cultural category (man or woman), based on their behaviour and social role. The latter is seen to derive from the former: a person's gender is an inevitable consequence of their sex [101].

Sociologists of gender—particularly ethnomethodologists, who study how individuals understand and reproduce the roles and constructs of society [32], and whose model of "doing gender" is now widespread within the social sciences[82]—have been particularly interested not just in how people model and gauge the gender of others, but how they *believe* it is modeled. As would be expected in a schema where gender derives from sex (indeed, the folk understanding often treats them as one and the same[49]), gender is seen as:

(1) *Binary*. There are two options: "man" and "woman" [88].
(2) *Immutable*. Once assigned a category, a person cannot alter their category [49].
(3) *Physiological*. Assignment is normatively handled on the basis of externally expressed physical characteristics, namely genitals[102]. In day-to-day interactions, dimorphic bone structure differences, hair patterns and other features act to distinguish "men" and "women" post-puberty [49].

This binary sex/binary gender view has long been understood to be inaccurate [42]. Even on a purely biological basis, research has shown a vast range of intersex conditions in which individuals do not match the common criterion for assignment to *either* sex [29]. Ethnomethodologists have found that while physiology is a component of how people infer the gender of others [61], its treatment as the sole element is very much *post hoc*: people make determinations about each other based on postures, dress, vocal cues, and then justify it with physiological cues after the fact [49].

Moreover, the assumption that sex dictates gender—in other words, that it mandates social roles, combinations of behaviours and traits and aspects of presentation and identity—fails to capture the existence of transgender (trans) people,[1] whose genders do not match their assigned sex. Trans people are contrasted with cisgender (cis) people, whose gender roles are congruent with their assigned sex.

Broadly speaking, trans people are "people whose gender identity or gender expression differs from expectations associated with the sex assigned to them at birth" [12]. Note that "differs" has a wide meaning: some individuals (trans men and trans women) fall within the gender binary but have a gender incongruent with the expected gender of their assigned sex. Others have non-binary genders, which do not fit neatly into either binary option; yet others fluctuate between genders (genderfluid) or have no gender at all (agender) [79]. Trans existences are not explained or captured by the traditional view, and so gender theorists have attempted to come up with new models [53, 65]. While they vary in their ideas of *precisely* what gender is or how it comes about, they generally agree that gender is not immutable, binary or tied inherently to physiology. I will call operationalisations with these features "trans-inclusive views".

Despite the clear limitations of the traditional view in capturing both the biological and cultural range of humanity, it has long been the standard way gender is operationalised in much of Western society, and as a consequence has been codified into everything from language to the design of physical spaces. Research discusses essential "male" and "female" differences [98]; architects make bathrooms for "men" and "women" only, with design features based on assumed physiological differences; medical training and processes are designed only for situations where someone's gender consistently matches their anatomy and resulting medical needs [25]; even clothing design assumes a bimodal range of presentations, coupled with a bimodal range of physical dimensions [59]. Trans people are simply not considered (i.e. erased) in much of public life and contemporary understandings of the world [68].

This erasure is a foundational component of the discrimination trans people face. If systems are not designed to include trans people, inclusion becomes an active struggle: individuals must actively fight to be included in things as basic as medical systems[75], legal systems[35] or even bathrooms[11, 91]. This creates space for widespread explicit discrimination [68], which has (in, for example, the United States) resulted in widespread employment, housing and criminal justice inequalities [36], increased vulnerability to intimate partner abuse and[35], particularly for trans people of colour, increased vulnerability to potentially fatal state violence[45].

2.2 Gender in HCI

Human-Computer Interaction (HCI) has long-studied the role gender plays in system usability and access[19, 84], to the point of having a dedicated subfield ("Gender HCI")[10].

But this is not to say that it has avoided erasing trans concerns—to the contrary, Gender HCI *itself* has been noted as largely rendering gender as fixed[16]. Exploring HCI's models of gender in 2011, Rode found that they "assume gender is stagnant and based on physiology" [83], noting a need to "move past binary gender in order to allow a flexible discussion of gender and technology". The same year, Kannabiran critiqued the "pronounced silence about the existence of non-binary genders" within the field—a critique validated by Schlesinger *et al.*, who showed in 2017 that only three papers in the history of CHI proceedings focused on trans users, and none centered non-binary

[1] In line with other research in this field [38], I will simply say trans for the rest of this paper.

users specifically [48, 89]. While trans users and trans concerns exist, this is not reflected in the literature.

As with wider society, this erasure has led to direct harm. Bivens & Haimson show how Facebook's approach to gender settings and decisions around photo tagging and recognition harms users going through transition [14]; Kannabiran's critical analysis found that users with non-binary identities are shut out of entire swathes of the platform [48]. Given the consistent critiques of HCI for its silence on trans issues, it seems inevitable that (unless the field has changed), it still contains the potential to cause yet more harm.

2.3 Automatic Gender Recognition

These concerns—trans-exclusive gender operationalisations in society, and similar operationalisations in HCI—intersect in Automatic Gender Recognition (AGR). AGR purports to allow the automatic, computational identification of a person's gender from photographs or videos. Implementations first isolate the person within a photograph: some use geometric structure[96], while others rely on skin texture[13], and yet others depend on 3D modelling [41]. The resulting image can then be subject to "gender recognition" which—while experimentally undertaken using gait[107] or overall body shape [80] —is usually based on the person's face [69].

A regular feature of facial recognition literature since 1990 [34], AGR research has many suggested applications. Some are purely computational; in theory, if one has an algorithm for recognising specific faces from a database, including a gender component improves speed by dramatically reducing the search space once gender has been detected [37]. Others include gendered access control in spaces such as bathrooms or changing rooms, with the AGR implementation triggering an alert should someone of the "wrong" gender enter a space[37], or gendered advertising and user interfaces, in which adverts or applications could (upon detecting a particular person's gender) alter their presentation to be more appealling to stereotypical members of that gender [55].

AGR is still in its infancy: there are some early real-world deployments of it for demographic analytics[24] and gendered shopping recommendations[33], but many commercial applications are little more then prototypes [77]. This has not stopped it from being used within the field of HCI, where AGR has been relied upon to understand (amongst other things) gendered dynamics in social media [3, 21, 62], or stopped the National Institute of Standards and Technology (NIST) from actively assessing and reporting on it due to its increasing motivation, potential and utility in commercial applications [70].

Precisely *why* this technology is necessary for, say, bathroom access control is not clear: most AGR papers do not dedicate any time to discussing the purported problem this technology is a solution to. The only clue comes from the NIST report mentioned above, which (while discussing the possible costs of false-positives in access control) states that: "the cost of falsely classifying a male as a female (i.e., the false female rate) could result in allowing suspicious or threatening activity to be conducted" [70], a statement disturbingly similar to the claims and justifications made by advocates of anti-trans "bathroom bills" [81].

2.4 Studying AGR

Given the systemic nature of trans erasure and harm in both public life and HCI, the worrying resonances of how AGR researchers discuss its potential uses, and its current and pending deployments in research and social life, it is worth critically exploring AGR and its uses. Doing so touches on a nascent CSCW subfield that studies how algorithms are designed, deployed and understood [100]. Prior works have looked at user understandings of algorithmic decisionmaking, and how this is influenced by the opacity (or transparency) of an algorithm's implementation [28, 52], and the costs and risks of assuming that an algorithm designed based on a constrained population or framing of a problem will adequately represent reality [92].

AGR is not an unstudied space; Hamidi *et al.* [40], motivated by similar concerns to mine, have already qualitatively interviewed trans people about their reaction to the prospect of AGR's deployment. But my focus is somewhat different; rather than user responses to hypothetical deployment of a hypothetically designed system, I am interested in how AGR research and HCI consumers of the technology *as it is designed* wrestle with the concept of gender, what the potential consequences of that are for trans people, and how its use can inform our understanding about the way HCI *overall* considers gender in its research and design work.

These are not abstract problems to me: I am a trans HCI researcher in *both* senses of the phrase, something I disclose because doing so helps situate my research: this is the work of someone who knows all-too-intimately the costs of trans-erasing systems[2]. At the same time, trans people are not a monolith: we have many identities and models of gender and self, and I can only speak to one experience directly. This paper should not be taken to be "the trans view" of this technology, any more than this paper (situated as it is within the Western view of gender) should be taken as universally applicable between cultures.

3 METHODOLOGY

3.1 AGR data and methods

To answer the first research question (*How does Automated Gender Recognition research operationalise gender, and what are the possible consequences of this should it be widely deployed?*), I chose to examine the field's artifacts—specifically, its papers. Works were gathered from the top 7 pattern recognition publication venues; as I am not a researcher in that field, I relied on venues' overall H-index, as determined by Scimego (see Table 1) [90]. By focusing on venues with high H-indexes (and consequently, high rates of citation), my hope was that I would also capture most of the high-citation *papers*, and consequently be able to get a good idea of the field's overall norms around gender in an efficient manner.

AGR paper sources

Venue	*H-index*	*Papers*
IEEE Transactions on Pattern Analysis and Machine Intelligence	288	8
Proc. IEEE Conference on Computer Vision and Pattern Recognition	192	7
Pattern Recognition	160	19
International Journal of Computer Vision	160	4
Proc. IEEE International Conference on Computer Vision	138	5
Journal of the Optical Society of America A	132	1
Pattern Recognition Letters	122	14

Table 1. Publication venues of the Automatic Gender Recognition (AGR) papers used in this content analysis.

I examined the entire archive of each venue, looking for facial recognition papers that either specifically worked on AGR technologies ("gender-focused" papers), or used AGR as a "test scenario" to benchmark a more general methodological contribution ("non-gender-focused" papers). Papers were identified and classified by reading the abstracts. As examples: Lapedriza et al.'s "Gender

[2]I confess to an ulterior motive as well: when I joined this field I only knew of *one* other trans researcher. Seeing a paper explicity affirm that we exist would have meant a lot to me: my hope is that it might help someone who is, reading this, where I was then.

Recognition in Non Controlled Environments" [51], which experiments with ways to approach AGR which work outside of laboratory settings, would be considered gender-focused. Yang et al.'s "Beyond Sparsity" [104], which argues for the use of a particular machine learning technique in facial recognition as a field, and uses the problem of gender recognition as one of a series of test cases to demonstrate that technique's utility, would be considered non-gender-focused.

Each paper that met the criteria for inclusion was then subjected to a content analysis, in which the text and graphs of the papers were coded so that I could understand how each work operationalised gender. In particular I looked for implicit or explicit reliance on each of the components of the traditional view of gender discussed previously: whether it treated gender as *binary* (consisting of only two categories), *immutable* (impossible to change once defined) and *physiological* (rooted in external, biological features). Examples of each, and of my process for determining whether a paper relied on a component, can be seen in the *Findings* section.

Content analysis was chosen as a methodological approach because it sits at the intersection of CSCW and traditional work on understanding encoded gender roles and assumptions. Within CSCW, content analyses have examined self-disclosure on social media [4, 74], teamwork on Wikipedia [66], and online privacy norms (with gender appearing as a facet of the analysis)[30]. In gender studies, the technique is ubiquitous: as early as 1993 researchers were complaining that "journals seem glutted with gender studies of contents" [47], and *Sex Roles* published an entire special issue on the method in 2010 [86]. It has commonly been applied to papers specifically in order to understand the trans-exclusivity of academic fields, and the consequences of that exclusivity [5, 15, 87, 93].

In total I found 58 papers, with publication years ranging from 1995 to 2017 (the last year of data I included in my analysis). Despite the wide range of publication years, it is clear from the distribution that AGR has been a far greater focus of attention recently than historically: the 10 years from 1997 to 2007 saw 4 papers published, while 2007 to 2017 saw 52. This dataset certainly has biases and limitations due to the focus on venues with high H-indexes: it is possible that, in my effort to capture high-citation papers that could be expected to set field norms, I have unintentionally *missed* high-citation papers that happened to be published in more subfield-specific, and so less widely cited, venues. Having conducted my analysis, I suspect that this is a low probability—as shown in the Findings section, the way the papers I selected operationalise gender are so consistent as to suggest little likely deviation from my results in the field as a whole. In other words, it is improbable that such papers would be *so* different as to change my overall results.

3.2 HCI data and methods

To answer my second question (*How does HCI research interacting with AGR operationalise gender and contextualise any gendered assumptions of AGR software?*) I looked at works from ICWSM, HT and the SIGCHI conferences and journals. Papers were included if they used an implementation of AGR as part of their methodology. In total I found 12 papers, spread over 4 venues, that met these criteria (Table 2).

All papers were from 2014 or later, with 5 appearing in 2017 (the last year I examined), suggesting that AGR is a novel (but increasingly utilised) method in the field. To examine how these papers operationalised gender, I replicated the analysis of gender used in examining AGR papers (looking for signs of binary, immutable or physiological assumptions). I then performed a content analysis to understand how the paper discussed AGR—whether it disputed the technology's model of gender, recognised any limitations in that model, or undertook efforts to actively test whether the model matched reality—and whether the paper's model and the model used by AGR were compatible.

	HCI paper sources	
Venue	*SIGCHI venue?*	*Papers*
Conference on Human Factors in Computing	Yes	4
International Conference on the Web and Social Media	No	4
Conference on Hypertext and Social Media	No	3
Conference on Recommender Systems	Yes	1

Table 2. Publication venues of the Automatic Gender Recognition (AGR) papers used in this content analysis.

4 FINDINGS

4.1 R1: Gender Operationalisation in AGR

For R1, I read each paper and evaluated whether it assumed gender was binary, immutable and/or physiological. The results were aggregated both overall, and split between gender-focused papers (papers explicitly developing AGR) and non-gender-focused papers (papers using AGR to test a more general recognition algorithm). Results can be seen in Table 3.

		Analysis results	
Paper type	*Binary*	*Immutable*	*Physiological*
Focused on gender	92.9%	71.4%	82.1%
No gender focus	96.7%	73.3%	40%
Overall	94.8%	72.4%	60.3%

Table 3. The results of my content analysis of 58 AGR papers. Each column contains the percentage of papers that explicitly or implicitly relied on one particular component of the traditional view of gender. Papers are also divided by whether they were focused particularly on gender, or simply used AGR as a test scenario for a more general facial recognition contribution.

My content analysis found a remarkably consistent operationalisation of gender within AGR research. Almost every paper with a focus on gender, and many of those without, treated gender in a way aligned with the traditional view. The few papers which did not rely exclusively on this view are largely those which did not discuss their model of gender, or essentialised (that is: treated as a component of what gender "is", or what gender "should be") elements of external appearance *other* than physiology.

4.1.1 Binary. Papers treated gender as binary **94.8%** of the time (92.9% in gender-focused papers, and 96.7% in non). By this I mean that they assumed that gender was a concept containing two, and only two, categories. An example of this being done explicitly would be:

> *"Gender classification is the binary classification problem of deciding whether a given face image contains a picture of a man or of a woman."* [8]

My reason for classifying that paper as featuring a binary treatment of gender is fairly obvious; it says gender is a binary problem. Some are slightly more implicit:

> *"For gender classification, we manually labelled gender information for the data set of Susskind et al. (2010). Thus, the data in the gender classification experiment are exactly the same, except for labels that are changed from expressions to gender (male/female)."* [109]

Papers near-uniformly fell into one of these two types of statement; either they would explicitly come out and claim that gender was a binary classification problem, or contained two categories, or they would commonly describe their dataset labels as only containing male/female or man/woman options, without any mention that this might be missing something.

4.1.2 Immutability. Immutability—the idea of gender as an unchanging status that cannot be altered post-assignment—was less discussed in AGR papers, which is reflected in the results: **72.4%** (71.4% gender-focused, 73.3% not) operated under the premise that gender was immutable. The lower percentage of papers falling into this classification comes due to many papers simply not discussing it, rather than any acceptance of gender as a state that can be altered. An example of where it is (implicitly) discussed can be found in one particular paper during its discussion of errors in its algorithm's output:

> *"The fourth image (blown up in 6e) is particularly interesting. This was tagged as female but we suspect it is a man in a wig!"* [1]

This is a case of binary thinking as well, but the important thing is that the researchers' reaction to seeing an unexpected combination of "female" clothing and "male" physiology: the person must be a man. They do not consider the possibility that the person is trans, or gender non-conforming: one cannot change their gender. Similarly:

> *"...a fully automatic human profile recognition system in which convolutional neural networks are learned for 3 important biological traits, i.e., human gender, age, and race." [106]*

Now, anthropologists agree that race is a socially constructed [44, 73], rather than purely biological, phenomenon. But if we are going to treat race as biological, and ignore the variations in what race "is" in different social contexts, the placement of race and age together creates a commonality: traits that cannot be changed by their holder. If that is the meaning of 'biological' in Yang *et al.*'s work, it further implies that gender is seen in the same way—it cannot be voluntarily altered. It is, for all intents and purposes, immutable.

4.1.3 Physiological. Establishing that a paper constructed gender with a *physiological* component—where the externally-visible structure of a person's body is the crucial factor in their gender—consisted largely of looking for references to physical traits being a factor in distinguishing gender:

> *"both networks are sensitive to the salient regions of the face: eyes, eyebrows, nose and mouth. The gender [network] is more sensitive to the centre of the mouth and to the periocular region."* [7]

Sometimes papers were even more explicit about their premise and/or approach being physiological:

> *"Here, our aim is to capture the morphological sexual differences between male and female faces by comparing their shape differences to a defined face template. We assume that such differences change with the face gender."* [103]

A comment of one or sometimes both of these types appeared regularly in the literature—papers would frequently reference the physiological features they were trying to use, or explicitly identify sexual dimorphism as their intended mechanism of distinction—yet only **60.3%** of papers (82.1% of those with a gender focus, 40% without) used a physiological model. Why the discrepancy?

In many cases, papers were simply too general to spend much time talking about gender; many non-gender-focused papers were seeking to make a field-wide contribution, and so focused discussion on problems common to all or most facial recognition tasks. Papers discussed new techniques to extract data from images in a more useful way [105], or new algorithms for facial recognition systems to "learn" with [76]. Gender was not discussed because the algorithm's *specific* efficacy for gender was not of interest.

In other instances I found that while papers did not essentialise physiology, they did essentialise other elements of personal appearance and expression:

> *"Attribute correlation can be either positive or negative. For example, a person with goatee and mustache is more likely to be a male, and is less likely to wear lipstick."* [103]

So while *physiology* was not the essentialised component of a person, essentialism was still occurring. Overall, it seems clear that AGR presents gender as a binary, often-immutable and either physiologically or otherwise essentialised concept.

4.2 R2: Gender and Gender Recognition in HCI

4.2.1 HCI gender models. The 12 HCI papers were subject to a similar analysis, looking at both their gender models and the way they described and used AGR. Most of them used common commercial implementations, while two developed their own machine learning system using a standard AGR dataset. Almost all of the papers were examining behaviour on social networks.

In their own framings of gender, none of the papers mentioned trans people or gave any indication that gender was more complex than the traditional model: many are very transparent about using a binary operationalisation:

> *"To overcome such limitation, in this work, we use the profile picture's URLs of all users in our dataset and use the Face++ API...to infer the gender (i.e., male or female), race (limited to Asian, Black, and White)."* [78]

By specifying that three racial categories are "limited", and making no such specification for gender, the paper makes clear that gender is considered a binary. Other papers treated gender and sex as interchangeable:

> *"We used the face recognition software Face++ to estimate the gender and age of users...Face++ uses computer vision and data mining techniques...to generate estimates of age and sex of individuals from their pictures."* [108]

Despite these hints of how gender is operationalised, it is worth pointing out that *none* of the papers made their operationalisation explicit. Every hint of what gender meant in the context of a paper was similar to the examples above; implicit, indirect, necessitating inference. This makes it difficult to say that there *is* a theory of gender behind this research, let alone what it is, but the clues provided strongly suggest a traditional model.

4.2.2 HCI treatment of AGR. None of the papers described AGR's model of gender at all, and consequently none either expressed discontent with or noted any limitations of said model. This is not to say that their AGR systems were not tested; papers regularly highlighted the reliability of their system to show that the paper's methods were robust:

> *"[the AGR software] itself re-turns the confidence levels for the inferred gender and race attributes, and it returns an error range for inferred age. In our data, the average confidence level reported by [the AGR software] is 95.22 ± 0.015% for gender and 85.97 ± 0.024% for race, with a confidence interval of 95%."* [62]

Note that this test of reliability does not examine whether the operationalisation maps to the users, merely that results are consistent with what its operationalisation dictates they *should* be. One attempt at testing against user-defined gender did appear in another paper:

> *"As an easy method to validate how accurate [the AGR software] is in inferring demographic information, we look at the profile description (i.e. bio) of Twitter users. We tag users who describe themselves as '(boy|guy|husband|father|dad|dude)' as 'Male' and (girl|wife|mother|mom) as 'Female'...concerning gender, we find that of the 2,433 users with one of the female indicator terms in their bios 82% are recognized as female by [the AGR software]. Male has an even higher detection rate-86% of 2,033 males who use one of the male indicator terms are detected as male by [the AGR software]."* [3]

Here the authors *have* treated gender as self-described, and as tied into social roles—but they still see gender as a binary, and do not question (or mention) the physiological rather than social basis of AGR's technologies. Even relying on that binary, the software still contradicted user self-descriptions 15-20% of the time[3]. These tests are also sometimes taken as legitimising AGR usage (and so AGR's operationalisation of gender):

> *"the authors used Twitter data to analyze the difference between men and women...the demographic characteristics of each user were obtained using Face++ and the Twitter user's profile picture...the researchers show that the strategy of getting demographic data from Face++ is reliable and provides accurate demographic information for gender and race, encouraging the application of this strategy in other recent efforts. We use a similar strategy to gather demographic information."* [78]

While many papers contained "limitations" sections or otherwise noted caveats, none of them discussed either the gender constraints of AGR or the often-identical constraints of the paper using it. To summarise, then: papers within HCI that use this technology do not note any limitations in its operationalisation of gender, do not test that operationalisation when assessing accuracy, and often replicate that operationalisation in their work without explicitly discussing it.

5 DISCUSSION

I have presented findings from an analysis of papers in the field of Automatic Gender Recognition (AGR), seeking to understand how researchers operationalise gender and provide examples of what the likely consequences of that operationalisation are for trans people (RQ1). In addition, I explored how papers in HCI which rely on this technology frame it, and how they operationalise gender themselves (RQ2).

With AGR (Section 4.1) I found that, where papers make their model implicitly or explicitly known, research near-uniformly relies on the traditional model of gender; that it follows a binary, that it is physiologically based, and that it is immutable. Almost every paper in my sample explicitly operationalised gender as binary, and an overwhelming majority saw it as immutable. In both cases, those papers that did *not* clearly operationalise gender in these ways did not mention their operationalisation at all. A clear majority of papers also saw gender as being rooted in physiological differences. The remainder either did not mention it, or suggested that gender could consistently be inferred from an individual's overall appearance and presentation.

HCI papers (Section 4.2) were never explicit about how they modeled gender, despite intending to measure it; those hints that are given suggest that work is relying on the traditional, trans-exclusive view of gender (Section 2.1). Papers did not report on any aspect of AGR's model of gender, even in papers that explicitly contained limitations sections. While there were some attempts to test or validate AGR results, they did not question the gender model being used.

5.1 Gendered spaces, gendered violence

Based on the findings from my first research question, AGR is particularly likely to misclassify (and so discriminate against) trans people. The presumption that gender is physiologically-rooted cuts against trans people overall (and undoubtedly some cis people) by essentialising the body as the source of gender. The presumption that gender is a binary additionally harms non-binary people, who by definition cannot be accurately classified. Both of these things are a problem when the technology is integrated with binary, gendered spaces—such as bathrooms.

As touched on in Section 1.1, the gendering and nature of bathrooms has complex implications for trans people. Because of the way gender is traditionally understood as a binary, there is frequently a complete absence of gender-neutral spaces for the purposes these rooms serve [97]. As a consequence, non-binary and agender people are often left having to pick between two incorrectly-gendered facilities. This causes a vast array of issues; the very need to choose denies their gender, causing substantial discomfort and embarassment [23], and leaves them (particularly if they are ambiguously-presenting) at risk of violence or threats from other, cis users of the facilities who perceive their presence as some kind of threat [26]. While trans men and trans women will usually (unlike people outside the binary) find a bathroom that matches their gender, they otherwise experience the same risks of hostility, rejection and assault [23, 26].

AGR's approach would make these problems worse. The most direct threat comes from how AGR papers propose handling ambiguously or "incorrectly" gendered subjects who enter a gendered space, and are registered as having a gender that does not match that space. The literature proposes having an "operator" (presumably some kind of facility security) alerted when this happens [20, 70]. What the operator then does is unclear. It could be checking the trans person's ID to see if their legal gender matches the gendering of the space—a substantial problem given how difficult it can be to *access* ID gender changes in many places [58]. It could be forcing the person to use the bathroom matching the gender they were assigned at birth, or to leave, the first of which is (as discussed) profoundly unpleasant and sometimes dangerous and the second of which would functionally exclude trans people from public life [43].

Most dangerously, an operator could call the police. This is not an abstract possibility; having the police called for trying to use the bathroom is a thing that already happens to trans people [43], even in the absence of a system that automatically alerts an observer to one's presence. Given the ubiquity of police discrimination and violence against trans people [36, 63, 71, 94], this means AGR would (if implemented as intended) simply automate the possibility of violence. The situation is likely to be worst for trans feminine people of colour; not only are they already subject to a disproportionate amount of discrimination and violence from the police [45], AGR has *already* been shown to be inaccurate at classifying feminine people with dark skin [18]. The technology's erasure of trans people, mirroring societal erasure discussed in Section 2.2, thus risks reinforcing the pre-existing discrimination and sometimes-fatal harm trans people face in the world.

5.2 Misgendering

A more insidious form of harm is misgendering; referring to someone with gendered terms that do not match their gender. This often takes the form of using incorrect pronouns (referring to a woman as "he", say); or honorifics ("Mr"). and is a common daily experience of trans people [6]. To be misgendered, particularly when it is on an ongoing basis, is to reinforce the idea that society does not consider your gender "real". The experience of being misgendered by people has been shown to prime individuals for rejection [85], impact self-esteem and felt authenticity, and increase one's perception of being socially stigmatised [60].

The proposed use cases of AGR both enable and directly contain misgendering, particularly given that we now know (thanks to the findings from RQ1) it is highly likely to misclassify trans users. To expand on the bathroom example: if a trans person is misclassified and then intercepted by the operator or police, they risk falling victim to threats or violence. But even if these things do not occur, the mere fact of their interception is an implicit claim by the system (and then person sent to "validate" their gender) that what the trans person knows about their own gender is not true.

Other AGR use cases create the same problem. AGR papers propose billboards with gender recognition technology that, after evaluating a pedestrian walking nearby, "may choose to show ads of cars when a male is detected, or dresses in the case of females" [95].[3] This may seem relatively innoccuous, but a trans man who sees a billboard flicker to advertise dresses to him as he approaches is, even if he likes dresses, unlikely to feel particularly good about it. For the billboard to have done that under the design specifications quoted above, it must have concluded he was a woman. Other papers experiment with recommender systems that present a range of different items to purchase [56], which offers the opportunity for even more explicit (and jarring) automated misgenderings given the vast number of products labeled "for men" and "for women" in the world.

5.3 Reinforcing erasure

Because AGR treats gender as a binary and physiological phenomenon (see Section 4.2), there is the potential not just for active harm (misgendering or the enabling of violence) but also erasure; the perpetuation of a normative view that trans people do not exist as a population with needs. This occurs in part because the adoption of AGR constrains designers in their subsequent work.

To return to bathrooms: suppose an architect designs a new building containing these spaces, and as part of that design decides to use AGR for access control. In doing so—in making their design dependent to conforming to that technology's limitations—they encode implicitly the idea that these spaces are gendered, and that this gendering should divide humanity into two categories. As a result they are unlikely to go out of their way to include non-gendered spaces—uniformly-gendered spaces are the path of least intellectual resistance, because the alternative is designing a space that certain promised features simply do not "work" in. Non-binary and agender users will be excluded from consideration in the resulting design, and as a consequence of the AGR usage, could find themselves barred from access to every bathroom in the building.

Within HCI applications of this technology (namely social media research), erasure makes itself known through the populations that cannot be reliably studied. A social media study based on AGR is likely to misclassify trans people, which means that trans communities generally (and non-binary communities in particular) simply cannot be studied. In addition, more general analyses will presumably either misclassify many trans users or be unable to reach an unambiguous result, producing messy data or an incentive to remove those datapoints from the study. Using AGR as a method in research again constrains thinking around what model of gender to use in the analysis, methods and user considerations; if a HCI researcher's method only allows them to gather data for a binary model of gender, they are going to end up with research that presumes gender is a binary. It is noticeable that every HCI paper that built its methodology around AGR also otherwise modeled gender as a binary.

A more general issue is that while AGR looks for physiological clues to gender, social media researchers are most-often looking for social differences between genders, and assuming

[3] This sort of casually assumptive statement is everywhere in the literature. For example, several papers propose demographic analytics to count the number of women walking into retail establishments. None propose counting the number of men. Can men not shop?

that one aligns with the other. For trans and gender non-conforming people—whose physiology and appearance may vary widely, and whose social experiences are very different from those of cis/gender-conforming people of the same gender—they do not. Even when the algorithm accurately identifies a trans woman as a woman, there are many layers of detail specific to trans experiences and social contexts that such a shallow model of gender will miss.

5.4 Fixing AGR?

Some might ask whether the issues with AGR raised in Section 4.2 and earlier parts of the discussion could be treated merely as an unfortunate but necessary inaccuracy. Trans people are, after all, a very small segment of the population, and all algorithms have a certain error rate we would deem acceptable. But the problem is not that there are errors but the *context* of those errors. First, an error rate that disproportionately falls on one population is not just an error rate: it is discrimination. It is precisely what is meant by algorithmic injustice[72]. Second, trans people are overwhelmingly the target of societal error rates: as discussed in Section 2.1, the codification of the traditional view means that they are treated as an outlier in almost every environment. Replicating these issues in new spaces serves only to perpetuate them, when technology is allegedly meant to *improve* the quality of our existences.

A person could also ask whether there is any way to design a trans-inclusive gender recognition system; making sure that there are trans people in the dataset, for example. People have tried to do this multiple times [50, 54]. Strictly speaking, they succeeded; they produced AGR systems that included trans people. But AGR systems are still categorical—how would one design a system that included all possible non-binary genders? —and still assuming a strong overall relationship between physiology and gender.

But whether or not AGR can be made to work in a *technically* trans-inclusive way does not answer the question of whether that is meaningful. Whatever approach (physiology, clothing, hair length...) an AGR system takes for discriminating between genders, however many trans people the dataset includes, the technology is fundamentally premised on the idea that gender is something *assigned*. Yet to be trans—to be of a gender that runs contrariwise to that which society assumed of you—means to stand as testament to the idea that it is self-knowledge, not external assignation, that has primacy in defining gender. Put simply, a trans-inclusive system for non-consensually defining someone's gender is a contradiction in terms.

5.5 Shining light on HCI's approach to gender

It should be clear at this point that I believe the deployment or use of AGR poses a substantial risk to trans people. Further than that, however, I believe that the swift acceptance of this technology provides an important lens on to how HCI perceives gender, and whose exclusion it considers worth tolerating. My findings in Section 4.2 show that AGR has been repeatedly and increasingly used in HCI, without any note in any publication about the constraints and risks that come with it.

On its own, this could be considered simply a fluke; after all, many parts of HCI engage in many different types of work, social media research being only one. The problem is that it adds to a large amount of existing work discussed earlier that shows trans concerns are rarely mentioned in HCI publications [89], and gender is frequently operationalised in a trans-exclusive fashion [83]. This is not specific to social network research: this is a phenomenon that has been historically found in all corners of HCI. It seems unlikely that everywhere *but* social media research has resolved it.

If concerns around trans-inclusivity were new, this issue would be broadly understandable; it takes time to adapt. But they are not. It has been seven years since Rode called for "nuance to avoid essentialism and binary treatments of gender" [83], and just as long since Kannabiran critiqued the field's silence on non-binary users [48]. But a small handful of exceptions aside [27, 31], few papers even mention trans existences, despite trans presences in many of the use cases and situations that motivate the field's work. This does not just lead to an absence of work on trans-specific issues but, as demonstrated by AGR, the development of technologies that have the potential to cause explicit harm.

My only explanations are that much of the HCI community does not understand the cost of trans-exclusive systems—the psychological and often physical harm associated with running counter to how society considers gender to function—and that, while several works have advocated trans-inclusive methodologies, there are few examples of their use or of what adaptions would need to be made. It is for these reasons that I have focused on AGR—a technology that is actively being deployed into wider society—and will (in my design recommendations) provide examples of trans-inclusive research that can be looked to as examples.

A deeper problem is that, as discussed in Section 2.2, HCI research tends to use a traditional view of gender by default, even within Gender HCI, and rarely explicitly defines what view it is using (shown in my second set of findings). These are immovable stumbling blocks to making progress on trans inclusion, since the traditional view fundamentally excludes trans people. HCI cannot resolve this longstanding problem if researchers do not ensure they start their research with an operationalisation of gender that, at the very minimum, recognises the existence of trans people and the falsehood of binary, physiologically-premised models.

6 DESIGN RECOMMENDATIONS

In Section 4.1 I have shown that gender is operationalised in a trans-exclusive way in AGR research; further, in Section 4.2, that this is not challenged or considered in HCI's use of the technology, and that HCI papers relying on AGR are simultaneously silent on trans existences. In my discussion I examined the likely consequences of this, ranging from limitations in the generalisability and equity of HCI research to the possibility of active, physical and psychological harm should AGR be built into the infrastructure of wider society. I have a number of recommendations as to how designers and researchers can learn from this, discussed below.

6.1 Avoid implementing AGR

My first finding (Section 4.1) was that AGR's premise is that gender is expressed (and so can be measured) in a binary, physiological and immutable form. This does not align with inclusive views of gender, or the reality of the social roles and contexts AGR is intended to be used in. This is not a consequence of a bad implementation, but a foundational and unavoidable element of the technology. As a result, in both research and non-research contexts it threatens to further harm an already marginalised population, fundamentally undermining their autonomy. There is, *ipso facto*, no way to make a technology premised on external inference of gender compatible with trans lives. Given the various way that continued usage would erase and put at risk trans people, designers and makers should quite simply avoid implementing or deploying AGR.

Instead, I suggest that designers working on gendered artifacts reflect on two questions. The first is whether the artifact has to be gendered, and if so, how to gender it in such a way that it recognises a wide range of people. As an example, consider gendered bathrooms (yet again). These spaces tend to codify an exclusive view of gender into the physical world and marginalise those who do not fit within it. A more inclusive approach to this kind of design problem would evaluate whether *non*-gendered spaces would better map to a wide range of users, or, if the spaces must be gendered, ensure that the design includes space for users whose genders fall outside the binary and recognise the challenges that trans men and women face in spaces that are gendered according to default, ciscentric expectations.

The second question is whether gender is genuinely the variable that best serves what a designer is trying to achieve. One example is AGR papers' proposal to use inferred gender to inform what products a user is advertised, with the assumption that gender can be used as a proxy for the more precise values that inform purchasing decisions. But those values are often imputable as well, negating the need to infer and use gender, and advertisers have already begun to move from demographic user segmentation to behaviour or interest-based approaches[22]. Although such a shift does not *entirely* fix gendered assumptions in this kind of tracking, it certainly helps, particularly given entirely legitimate trans concerns about the consequences of gendered surveillance.[14] The same can be done in many other spaces: looking directly for the value gender was intended to approximate, rather than something as complex and fluid as gender itself.

6.2 Examine gender with inclusive methods

The limitations of AGR demonstrated in our first finding make it difficult to see how it can be used in HCI without narrowing the frame of research and severely damaging our ability to be sensitive to user autonomy.

Social network researchers (or anyone else in HCI intending to use AGR for large-scale research on gender) should instead aim to rely on *self-disclosed* gender information, noting the restrictions some platforms may put on the range of options available. When gender as a demographic value is not encoded into the platform being studied, researchers could rely on commonalities in how users disclose and describe their gender, something further discussed in the "Future Work" section.

Researchers could also opt for a more mixed-methods approach in which they explicitly ask users to self-describe their gender. HCI already contains examples of this being done[39], and discussion of how to do it in a way respectful of non-binary genders [46]. Both approaches avoid the restrictive nature of AGR, and in the process reinforce the autonomy and primacy of individuals in defining their own personhood.

6.3 Frame gender explicitly and with trans-inclusivity

My second finding and the subsequent discussion touched on the much wider question of how the field of HCI considers gender. I showed that researchers' interactions with AGR have perpetuated the technology's trans-exclusive framing and failed to consider or discuss its limitations. I tied this to a wider set of problems in how HCI operationalises and discusses gender, and the dearth of trans-inclusive works—problems that have been widely discussed and yet are evidently little closer to being resolved.

Researchers should not be relying on a model of gender that excludes trans and gender non-conforming people from a project's framing before work has even begun—certainly not as a default position. Instead, it is imperative that we (as individuals and as a community) adopt new, more inclusive models of gender that explicitly reject the idea of gender as a purely binary, immutable and/or physiological construct. The utility of a particular model will likely vary depending on what aspect of gender a particular researcher is trying to study, but I would point to Rode's paper (discussed earlier)[83] as a source of varied and inclusive models to use, and Ahmed's "Trans Competent Interaction Design", and its discussion of user-centered processes for deciding on the adoption of said models [2].

In addition, researchers should explicitly discuss how gender is being operationalised in a particular work. Part of the problem here is that while quite a few papers talk about gender, few of them discuss what they think it *is*. As a consequence, researchers are not prompted to reconsider their framing while writing, and reviewers and readers struggle to directly confront any disconnect between a work's operationalisation of gender and the population affected.

If we are frank about what we mean by "gender", and if we (as Schlesinger *et al.* advised) "explicitly [report study] limitations in terms of identity categories" [89], we make it easier to contextualise our work, reflect on the consequences of our framing, and try to do better as a result. This is not a particularly large change—I am asking that when working with gender, researchers treat "gender" as a term of art and specify what they see it as in the context of their work. In other words, I am asking that researchers define their terms.

6.4 Make resources for a gender-aware HCI

Many of my recommendations start "researchers should", and I firmly believe that the onus is on the HCI research community to change how it evaluates gender the course of its work. At the same time, I recognise this is not a small task: change requires the resources and opportunities to reflect on the *status quo*, understand its limitations, and learn about and discuss better alternatives. Gender is infrequently discussed directly in a computing context, and even more infrequently unpacked. Reflecting on similar issues in 2014, Breslin & Wadhwa recommended that HCI-related degree programmes include a "Gender 101" course:

> *This ideally would be a whole course dedicated to exploring various gender theories, how gender relates to society, culture, economics, politics, language, and technology, and allowing students the scope and time to reflexively explore and understand the complexities of gender. However, at the very least students should be offered a class that addresses gender vocabulary (gender versus sex versus sexuality, gender roles, gender identity, etc.), different theories on gender (essentialist versus performative, for example), and different gendered norms cross-culturally or cross-context"*[16]

I see no reason why tutorials on precisely these issues, aimed at researchers and practitioners, could not be developed. Just as SIGCHI holds summer camps or conference courses on gesture production [99], sketching [57], and smart matter [67], it could host (and directly request) events aimed at critically exploring the conceptualisation and use of gender. Breslin & Wadhwa have seperately developed a model curriculum which could be used as a basis for this work [17].

In parallel, SIGCHI should enable researchers with interest and expertise in gender to collaborate on static, shareable resources: best practices guidelines for researchers or reviewers, lesson plans, guides to particular aspects of gender and conversations around it. Conference courses are tremendously valuable for an in-depth discussion, but are difficult to scale and can only properly reach the people who attend. More static educational resources can be distributed more widely, and are more easily updated as our understanding of gender increases.

7 FUTURE WORK

7.1 Further AGR analysis

My findings in this paper open up a variety of directions for further research. A full algorithmic audit which seeks to test "top performing" models against visual images of trans people would give us a better idea of precisely how inaccurate AGR is in practice, and thus the scale of the issues the technology invokes. This audit should rely on data which is not only gender-diverse but *racially* diverse: as briefly discussed, existing research has already shown that these algorithms have reduced accuracy at identifying "darker-skinned [women]" [18], and since feminine trans people of colour are disproportionately the subjects of existing violence and discrimination [36], an audit would be most useful if it could speak to the intersections of race and trans identities. While I have approached this paper with an insider view of "trans identities" broadly-construed, I navigate the world as a white person, and so my thoughts on how gender and *race* intersect are very much from the perspective of an outsider.

It could also be valuable to dig deeper into how AGR's model of gender is operationalised, and whether it also appears in shared datasets, codebases and the perspectives of the researchers themselves. A trans-exclusionary operationalisation of gender is not unique to AGR and HCI —this area is something many fields are struggling with [98]. By working to understand how gender operationalisations come to be and the lifecycle of their codification into artefacts, we can begin work on targeted interventions that aim to make the consequences of poor operationalisations, and alternative operationalisations, clear to technological researchers.

7.2 Designing replacement methodologies

There are several existing respectful approaches to gauging gender, discussed earlier, but I recognise that it can sometimes be difficult to apply qualitative or mixed-methods approaches to quantitative research questions or spaces. Consequently, we should not only critique the use of AGR but also work to develop better replacements, aimed at the kind of spaces where researchers currently deploy the technology. These replacements should respect user autonomy in defining the self, and recognise a wide range of possible descriptors this autonomous definition could produce.

One way of doing this for social network research would be to rely on user-provided profile data for gender inference, rather than an algorithmic analysis. In the case of Twitter, this could consist of looking for cues in pronoun usage, the use of particularly common informal terms for various genders ("enby", for example, is a common informal self-descriptor for someone with a non-binary gender), or common trans-specific hashtags (see #GirlsLikeUs) [64]. Not only would this provide a user-centered alternative for gathering information on participant gender, it would also allow researchers to add depth and nuance to their analyses by being able to consider trans-specific network dynamics and the interplay between trans and cis participants.

8 CONCLUSION

Automated Gender Recognition purports to recognise gender—but my analysis of the way in which it operates shows that this is only true if one denies the role that self-knowledge plays in gender, and consequently, denies the existence of trans people. Using a more inclusive view of gender, one that recognises the primacy of self, we can see that the systems are not recognising gender (which would be impossible to reliably do by inference), but instead imposing their own views of gender on unwitting users and research subjects. Thirty years of research has not produced a single system that can genuinely recognise gender—it has only produced systems that try to assign it.

This makes the rapidly-spreading adoption of AGR by Human-Computer Interaction highly concerning. The technology is of increasing popularity within the field, yet researchers using it never report the trans-exclusivity or other assumptions about the construction of gender, and often follow precisely the same model, resulting in the erasure of trans people from the research that relies on it—erasure that has already created space for direct harm. The way this echoes longstanding concerns about trans-inclusivity in HCI, and an absence of nuance in how the field often conceptualises gender, suggests that little progress has been made despite regular nudges and critiques.

HCI urgently needs to do better, both generally, by applying the "hermeneutics of suspicion" to the tools, methods and theories we integrate [9], and specifically, through operationalising and understanding gender in a nuanced way. Without active work, these problems are unlikely to get better over time: HCI will continue to both directly and indirectly harm trans people.

9 ACKNOWLEDGEMENTS

Everybody Knows that first-year PhD students do not write solo-author papers—and because Everybody Knows that, nobody thought to tell me until I was seven drafts in. I am accordingly tremendously grateful to David Ribes, Daniela Rosner, Charlotte Lee, Jennifer Turns, Frances Hocutt, Morgan Klaus Scheuerman, Talia Stein and Margret Wander. On a personal basis, I would be remiss if I did not thank Fred Unwin, Christine Kershner, Tanya Ewing, and Rachel F. Sanchez for their ongoing care and support.

REFERENCES

[1] Jania Aghajanian, Jonathan Warrell, Simon J.D. Prince, Peng Li, Jennifer L. Rohn, and Buzz Baum. 2009. Patch-based within-object classification. *Proceedings of the IEEE International Conference on Computer Vision* (2009), 1125–1132. https://doi.org/10.1109/ICCV.2009.5459352

[2] Alex Ahmed. 2017. Trans Competent Interaction Design: A Qualitative Study on Voice, Identity, and Technology. *Interacting with Computers* 30, 1 (2017). https://doi.org/10.1093/iwc/iwx018

[3] Jisun An and Ingmar Weber. 2016. #greysanatomy vs. #yankees: Demographics and Hashtag Use on Twitter. *Proceedings of the Tenth International AAAI Conference on Web and Social Media* Icwsm (2016), 523–526. arXiv:1603.01973 https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13021/12779

[4] Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2017. Sensitive Self-disclosures, Responses, and Social Support on Instagram. In *ACM Conference on Computer Supported Cooperative Work*. ACM Press, New York, New York, USA, 1485–1500.

[5] Y Gavriel Ansara and Peter Hegarty. 2012. Cisgenderism in psychology: pathologising and misgendering children from 1999 to 2008. *Psychology & Sexuality* 3, 2 (May 2012), 137–160.

[6] Y. Gavriel Ansara and Peter Hegarty. 2014. Methodologies of misgendering: Recommendations for reducing cisgenderism in psychological research. *Feminism and Psychology* 24, 2 (2014), 259–270. https://doi.org/10.1177/0959353514526217

[7] Grigory Antipov, Moez Baccouche, Sid Ahmed Berrani, and Jean Luc Dugelay. 2017. Effective Training of Convolutional Neural Networks for Face-Based Gender and Age Prediction. *Pattern Recognition* 72 (2017), 15–26. https://doi.org/10.1016/j.patcog.2017.06.031

[8] Abbas Roayaei Ardakany, Mircea Nicolescu, and Monica Nicolescu. 2013. An Extended Local Binary Pattern for Gender Classification. *2013 IEEE International Symposium on Multimedia* (2013), 315–320. https://doi.org/10.1109/ISM.2013.61

[9] Jeffrey Bardzell and Shaowen Bardzell. 2015. Humanistic HCI. *Synthesis Lectures on Human-Centered Informatics* 8, 4 (2015), 1–185.

[10] L Beckwith and M Burnett. 2004. Gender: An Important Factor in End-User Programming Environments?. In *2004 IEEE Symposium on Visual Languages - Human Centric Computing*. IEEE, 107–114.

[11] Kyla Bender-Baird. 2016. Peeing under surveillance: bathrooms, gender policing, and hate violence. *Gender, Place and Culture* 23, 7 (2016), 983–988. https://doi.org/10.1080/0966369X.2015.1073699

[12] Marla Berg-Weger. 2016. *Social Work and Social Welfare: An Invitation*. Routledge.

[13] Francesco Bianconi, Fabrizio Smeraldi, Maryam Abdollahyan, and Perry Xiao. 2016. On the use of skin texture features for gender recognition: An experimental evaluation. *Image Processing Theory Tools and Applications (IPTA), 2016 6th International Conference on*, 1–6.

[14] Rena Bivens and Oliver L. Haimson. 2016. Baking Gender Into Social Media Design: How Platforms Shape Categories for Users and Advertisers. *Social Media + Society* 2, 4 (2016), 205630511667248. https://doi.org/10.1177/2056305116672486

[15] Markie L C Blumer, Mary S Green, Sarah J Knowles, and April Williams. 2012. Shedding Light on Thirteen Years of Darkness: Content Analysis of Articles Pertaining to Transgender Issues In Marriage/Couple and Family Therapy Journals. *Journal of Marital and Family Therapy* 38, 1 (June 2012), 244–256.

[16] Samantha Breslin and Bimlesh Wadhwa. 2014. Exploring Nuanced Gender Perspectives within the HCI Community. *Proceedings of the India HCI 2014 Conference on Human Computer Interaction - IHCI '14* (2014), 45–54. https://doi.org/10.1145/2676702.2676709

[17] Samantha Breslin and Bimlesh Wadhwa. 2015. Towards a Gender HCI Curriculum. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (2015), 1091–1096. https://doi.org/10.1145/2702613.2732923

[18] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research* 81 (2018), 1–15.

[19] Margaret Burnett, Anicia Peters, Charles Hill, and Noha Elarief. 2016. Finding Gender-Inclusiveness Software Issues with GenderMag: A Field Investigation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2586–2598. https://doi.org/10.1145/2858036.2858274

[20] Modesto Castrillón-Santana, Maria De Marsico, Michele Nappi, and Daniel Riccio. 2017. MEG: Texture operators for multi-expert gender classification. *Computer Vision and Image Understanding* 156 (2017), 4–18. https://doi.org/10.

1016/j.cviu.2016.09.004
[21] Xin Chen, Yu Wang, Eugene Agichtein, and Fusheng Wang. 2015. A Comparative Study of Demographic Attribute Inference in Twitter. *Ninth International AAAI Conference on Web and Social Media* (2015), 590–593. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10541
[22] Elizabeth F. Churchill. 2013. Putting the Person Back into Personalization. *interactions* 20, 5 (Sept. 2013), 12–15. https://doi.org/10.1145/2504847
[23] Catherine Connell. 2011. Transgender and Genderqueer Workers Negotiating "the Bathroom Question". In *Embodied Resistance: Challenging the Norms, Breaking the Rules*, Chris Bobel & Samantha Kwan (Ed.). Vanderbilt University Press, New York.
[24] Sing Tao Daily. 2017. Xindi Mall Extension Technology Application Human Face Recognition System. (Dec 2017). Retrieved April 3, 2018 from http://hd.stheadline.com/news/daily/fin/626442/
[25] Vivek Divan, Clifton Cortez, Marina Smelyanskaya, and Joanne Keatley. 2016. Transgender social inclusion and equality: A pivotal path to development. *Journal of the International AIDS Society* 19, Suppl 2 (2016), 1–6. https://doi.org/10.7448/IAS.19.3.20803
[26] Petra L. Doan. 2010. The tyranny of gendered spaces - reflections from beyond the gender dichotomy. *Gender, Place & Culture* 17, 5 (2010), 635–654. https://doi.org/10.1080/0966369X.2010.503121
[27] Daniel A. Epstein, Nicole B. Lee, Jennifer H. Kang, Elena Agapie, Jessica Schroeder, Laura R. Pina, James Fogarty, Julie A. Kientz, and Sean Munson. 2017. Examining Menstrual Tracking to Inform the Design of Personal Informatics Tools. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17* (2017), 6876–6888. https://doi.org/10.1145/3025453.3025635 arXiv:15334406
[28] Motahhare Eslami. 2017. Understanding and Designing Around Users' Interaction with Hidden Algorithms in Sociotechnical Systems. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17 Companion)*. ACM, New York, NY, USA, 57–60. https://doi.org/10.1145/3022198.3024947
[29] Anne Fausto-Sterling. 2000. *Sexing the Body: Gender Politics and the Construction of Sexuality*. Basic Books, New York, NY, USA.
[30] Casey Fiesler, Michaelanne Dye, Jessica L. Feuston, Chaya Hiruncharoenvate, C.J. Hutto, Shannon Morrison, Parisa Khanipour Roshan, Umashanthi Pavalanathan, Amy S. Bruckman, Munmun De Choudhury, and Eric Gilbert. 2017. What (or Who) Is Public?: Privacy Settings and Social Media Content Sharing. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 567–580. https://doi.org/10.1145/2998181.2998223
[31] Andrea Forte, Nazanin Andalibi, and Rachel Greenstadt. 2017. Privacy, anonymity, and perceived risk in open collaboration: a study of Tor users and wikipedians. *Proceedings of Computer Supported Cooperative Work and Social Computing* (2017), 1800–1811. https://doi.org/10.1145/2998181.2998273
[32] Harold Garfinkel. 1968. The Origins of the Term 'Ethnomethodology'. In *Proceedings of the Purdue Symposium on Ethnomethodology*. Purdue University, 5–11.
[33] Globe and Mail. 2011. Ten ways AI can help you manage your life. (2011). https://www.theglobeandmail.com/life/article-ten-ways-ai-can-help-you-manage-your-life/
[34] B. A. Golomb, D. T. Lawrence, and T. J. Sejnowski. 1990. SexNet: A Neural Network Identifies Sex from Human Faces. In *Proceedings of the 1990 Conference on Advances in Neural Information Processing Systems 3 (NIPS-3)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 572–577. http://dl.acm.org/citation.cfm?id=118850.118953
[35] Leigh Goodmark. 2013. Transgender people, intimate partner abuse, and the legal system. *Harvard Civil Rights-Civil Liberties Law Review* 48, 1 (Oct. 2013), 51–104.
[36] J M Grant, L a Mottet, J Tanis, J Harrison, J L Herman, and M Keisling. 2011. Injustice at Every Turn: A Report of the National Transgender Discrimination Survey. *Washington National Center for Transgender Equality and National Gay and Lesbian Task Force* 25 (2011), 2011. https://doi.org/10.1016/S0016-7878(90)80026-2
[37] Abdenour Hadid and Matti Pietikäinen. 2009. Combining appearance and motion for face and gender recognition from videos. *Pattern Recognition* 42, 11 (2009), 2818–2827. https://doi.org/10.1016/j.patcog.2009.02.011
[38] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2015. Disclosure, Stress, and Support During Gender Transition on Facebook. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15* (2015), 1176–1190. https://doi.org/10.1145/2675133.2675152
[39] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2016. Digital Footprints and Changing Networks During Online Identity Transitions. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (2016), 2895–2907. https://doi.org/10.1145/2858036.2858136
[40] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. 2018. Gender Recognition or Gender Reductionism?: The Social Implications of Embedded Gender Recognition Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 8, 13 pages. https://doi.org/10.1145/3173574.3173582

[41] Xia Han, Hassan Ugail, and Ian Palmer. 2009. Gender Classification Based on 3D Face Geometry Features Using SVM. In *International Conference on CyberWorlds*. IEEE, 114–118.
[42] Judah HaNasi. 189. Mishnah Bikkurim. In *Mishnah*. Chapter 4.
[43] Jody L Herman. 2013. Gendered restrooms and minority stress: The public regulation of gender and its impact on transgender people's lives. *Journal of Public Management and Social Policy* (2013), 65–80.
[44] Nicholas Hudson. 1996. From "Nation to "Race": The Origin of Racial Classification in Eighteenth-Century Thought. *Eighteenth-Century Studies* 29, 3 (April 1996), 247–264.
[45] Angela Irvine. 2015. You Can't Run from the Police: Developing a Feminist Criminology that Incorporates Black Transgender Women. *Southwest Law Review* 44 (2015), 553–561. https://doi.org/10.3366/ajicl.2011.0005 arXiv:arXiv:1011.1669v3
[46] Samantha Jaroszewski, Danielle Lottridge, Oliver L Haimson, and Katie Quehl. 2018. "Genderfluid" or "Attack Helicopter": Responsible HCI Practice with Non-Binary Gender Variation in Online Communities. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (2018).
[47] Lee Jolliffe. 1993. Yes! More Content Analyses! *Newspaper Research Journal* 14, 3-4 (1993), 93–98.
[48] Gopinaath Kannabiran. 2011. Themself: Critical Analysis of Gender in Facebook. (2011), 1–6. https://feministhciworkshop.files.wordpress.com/2010/11/gkannabiran{_}feministhci2011.pdf
[49] Suzanne J Kessler and Wendy McKenna. 1978. *Gender: An ethnomethodological approach*. University of Chicago Press.
[50] Vijay Kumar, R Raghavendra, Anoop Namboodiri, and Christoph Busch. 2016. Robust transgender face recognition: Approach based on appearance and therapy factors. In *IEEE International Conference on Identity, Security and Behavior Analysis*. IEEE, 1–7.
[51] Àgata Lapedriza, Manuel J Marín-Jiménez, and Jordi Vitrià. 2006. Gender Recognition in Non Controlled Environments. In *IEEE Conference on Pattern Recognition*. IEEE, 834–837.
[52] Min Kyung Lee and Su Baykal. 2017. Algorithmic Mediation in Group Decisions: Fairness Perceptions of Algorithmically Mediated vs. Discussion-Based Social Division. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1035–1048. https://doi.org/10.1145/2998181.2998230
[53] Jason Lim and Kath Browne. 2009. Senses of Gender. *Sociological Research Online* 14, 1 (2009), 1–14.
[54] Gayathri Mahalingam, Karl Ricanek, and A. Midori Albert. 2014. Investigating the periocular-based face recognition across gender transformation. *IEEE Transactions on Information Forensics and Security* 9, 12 (2014), 2180–2192. https://doi.org/10.1109/TIFS.2014.2361479
[55] Jordi Mansanet, Alberto Albiol, and Roberto Paredes. 2016. Local Deep Neural Networks for gender recognition. *Pattern Recognition Letters* 70 (2016), 80–86. https://doi.org/10.1016/j.patrec.2015.11.015
[56] Jennifer Marlow and Jason Wiese. 2017. Surveying User Reactions to Recommendations Based on Inferences Made by Face Detection Technology. *Proceedings of the Eleventh ACM Conference on Recommender Systems - RecSys '17* (2017), 269–273. https://doi.org/10.1145/3109859.3109875
[57] Nicolai Marquardt. 2017. Sketching user experiences: Hands-on course of sketching techniques for HCI research. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* Part F1276 (2017), 6–8. https://doi.org/10.1145/3027063.3027107
[58] James McGrath. 2009. Are You a Boy or a Girl? Show Me Your REAL ID. *Nevada Law Journal* 9, 4 (2009).
[59] Kim McKeage, Elizabeth Crosby, and Terri Rittenburg. 2017. Living in a Gender-Binary World: Implications for a Revised Model of Consumer Vulnerability. *Journal of Macromarketing* 38, 1 (Dec. 2017), 73–90.
[60] Kevin A. McLemore. 2015. Experiences with Misgendering: Identity Misclassification of Transgender Spectrum Individuals. *Self and Identity* 14, 1 (2015), 51–74. https://doi.org/10.1080/15298868.2014.950691
[61] James W Messerschmidt. 2009. "Doing Gender": The Impact and Future of a Salient Sociological Concept. *Gender & Society* 23, 1 (Feb. 2009), 85–88.
[62] Johnnatan Messias, Pantelis Vikatos, and Fabrício Benevenuto. 2017. White, Man, and Highly Followed: Gender and Race Inequalities in Twitter. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, New York, NY, USA, 266–274. https://doi.org/10.1145/3106426.3106472
[63] Toby Miles-Johnson. 2015. "They Don't Identify With Us": Perceptions of Police by Australian Transgender People. *International Journal of Transgenderism* 16, 3 (2015), 169–189. https://doi.org/10.1080/15532739.2015.1080647
[64] Janet Mock. 2018. Growing Up Trans: Sisterhood and Shelter at the Hetrick-Martin Institute. (2018). https://janetmock.com/2011/07/23/transgender-speech-hetrick-martin-nyc/
[65] Surya Monro. 2010. Towards a Sociology of Gender Diversity. In *Transgender Identities: Towards a Social Analysis of Gender Diversity*, Sally Hines and Tom Sanger (Eds.). Chapter 4.
[66] Jonathan T. Morgan, Michael Gilbert, David W. McDonald, and Mark Zachry. 2014. Editing Beyond Articles: Diversity & Dynamics of Teamwork in Open Collaborations. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 550–563.

https://doi.org/10.1145/2531602.2531654
[67] Stefanie Mueller, Sean Follmer, Thijs Roumen, and Parinya Punpongsanon. 2017. SCF Summer School. (jul 2017). Retrieved April 12, 2018 from http://scf-summerschool.com/
[68] Viviane Namaste. 2000. *Invisible Lives: The Erasure of Transsexual and Transgendered People.* University of Chicago Press.
[69] Choon Boon Ng, Yong Haur Tay, and Bok-Min Goi. 2012. Vision-based Human Gender Recognition - A Survey. *CoRR* (2012).
[70] Mei L Ngan and Patrick J Grother. 2015. Face Recognition Vendor Test (FRVT)-Performance of Automated Gender Classification Algorithms. *NIST Interagency/Internal Report (NISTIR)-8052* (2015).
[71] Andrea Nichols. 2010. Dance ponnaya, dance! police abuses against transgender sex workers in Sri Lanka. *Feminist Criminology* 5, 2 (2010), 195–222. https://doi.org/10.1177/1557085110366226
[72] Safiya Umoja Noble. 2018. Algorithms of Oppression. (June 2018), 1–188.
[73] Stephan Pamlié. 2008. Genomics, divination, "racecraft". *American Ethnologist* 34, 2 (Jan. 2008), 205–222.
[74] Jessica A Pater, Oliver L Haimson, Nazanin Andalibi, and Elizabeth D Mynatt. 2016. "Hunger Hurts but Starving Works:" Characterizing the Presentation of Eating Disorders Online. In *ACM Conference on Computer Supported Cooperative Work*. ACM Press, New York, New York, USA, 1183–1198.
[75] Tonia Poteat, Danielle German, and Deanna Kerrigan. 2013. Managing Uncertainty: A Grounded Theory of Stigma in Transgender Healthcare Encounters. *Social Science and Medicine* 84 (2013), 22–29. https://doi.org/10.1016/j.socscimed.2013.02.019 arXiv:15334406
[76] Oriol Pujol and David Masip. 2009. Geometry-based ensembles: Toward a structural characterization of the classification boundary. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 6 (2009), 1140–1146. https://doi.org/10.1109/TPAMI.2009.31
[77] Corinne Reichert. 2017. Intel Demos 5G Facial-Recognition Payment Technology. (dec 2017). Retrieved April 3, 2018 from http://www.zdnet.com/article/intel-demos-5g-facial-recognition-payment-technology/
[78] Julio C. S. Reis, Haewoon Kwak, Jisun An, Johnnatan Messias, and Fabricio Benevenuto. 2017. Demographics of News Sharing in the U.S. Twittersphere. *Proceedings of the 28th ACM Conference on Hypertext and Social Media* (2017), 195–204. https://doi.org/10.1145/3078714.3078734 arXiv:1705.03972
[79] Christina Richards, Walter Pierre Bouman, and Meg-John Barker. 2017. *Genderqueer and Non-Binary Genders.* Springer, London.
[80] M. Rioux, Chang Shu, and S. Wuhrer. 2009. Posture invariant gender classification for 3D human models. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops(CVPRW)*, Vol. 00. 33–38. https://doi.org/10.1109/CVPR.2009.5204295
[81] Samantha Janette Rippetoe. 2017. *House Bill 2 and the Myth of the Bathroom Predator: Exploring Gendered Assumptions in the Context of "Livable Lives" in Policy Making.* Master's thesis. Stanford, CA, USA. Advisor(s) Atchison, Robert J. AAT 10278703.
[82] Barbara J Risman. 2009. From Doing To Undoing: Gender as We Know It. *Gender & Society* 23, 1 (Feb. 2009), 81–84.
[83] Jennifer A. Rode. 2011. A Theoretical Agenda for Feminist HCI. *Interacting with Computers* 23, 5 (2011), 393–400. https://doi.org/10.1016/j.intcom.2011.04.005
[84] Els Rommes, Ellen Van Oost, and Nelly Oudshoorn. 1999. Gender in the Design of the Digital City of Amsterdam. *Information, Communication & Society* 2, 4 (1999), 476–495. https://doi.org/10.1080/136911899359510 arXiv:http://dx.doi.org/10.1080/136911899359510
[85] Brian A. Rood, Sari L. Reisner, Francisco I. Surace, Jae A. Puckett, Meredith R. Maroney, and David W. Pantalone. 2016. Expecting Rejection: Understanding the Minority Stress Experiences of Transgender and Gender-Nonconforming Individuals. *Transgender Health* 1, 1 (2016), 151–164. https://doi.org/10.1089/trgh.2016.0012
[86] Rena M Rudy, Lucy Popova, and Daniel G Linz. 2010. The Context of Current Content Analysis of Gender Roles: An Introduction to a Special Issue. *Sex Roles* 62, 11-12 (July 2010), 705–720.
[87] K Scherrer and M Woodford. 2014. Incorporating Content on Gay, Lesbian, Bisexual, Transgender, and Queer Issues in Leading Social Work Journals. *Social Work Research* 37, 4 (Jan. 2014), 423–431.
[88] Kristen Schilt and Laurel Westbrook. 2009. Doing Gender, Doing Heteronormativity. *Gender & Society* 23, 4 (July 2009), 440–464.
[89] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. 2017. Intersectional HCI: Engaging Identity Through Gender, Race, and Class. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17).* ACM, New York, NY, USA, 5412–5427. https://doi.org/10.1145/3025453.3025766
[90] Scimago. 2017. Journal Rankings on Computer Vision and Pattern Recognition. Website. (14 November 2017). http://www.scimagojr.com/journalrank.php?category=1707&order=h&ord=desc
[91] Kristie L. Seelman. 2014. Transgender Individuals' Access to College Housing and Bathrooms: Findings from the National Transgender Discrimination Survey. *Journal of Gay and Lesbian Social Services* 26, 2 (2014), 186–206.

https://doi.org/10.1080/10538720.2014.891091
[92] Shilad Sen, Margaret E Giesel, Rebecca Gold, Benjamin Hillmann, Matt Lesicko, Samuel Naden, Jesse Russell, Zixiao Ken Wang, and Brent Hecht. 2015. Turkers, Scholars, "Arafat" and "Peace". In *ACM Conference on Computer Supported Cooperative Work*. ACM Press, New York, New York, USA, 826–838.
[93] Anneliese Singh and Kimber Shelton. 2011. A content analysis of LGBTQ qualitative research in counseling: A ten-year review. *Journal of Counseling and Development* 89, 2 (April 2011), 217–226.
[94] Rebecca L. Stotzer. 2014. Law Enforcement and Criminal Justice Personnel Interactions with Transgender People in the United States: A Literature Review. *Aggression and Violent Behavior* 19, 3 (2014), 263–277. https://doi.org/10.1016/j.avb.2014.04.012
[95] Yunlian Sun, Man Zhang, Zhenan Sun, and Tieniu Tan. 2017. Demographic Analysis from Biometric Data: Achievements, Challenges, and New Frontiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* X, X (2017), 1–1. https://doi.org/10.1109/TPAMI.2017.2669035
[96] Zehang Sun, George Bebis, Xiaojing Yuan, and Sushil J Louis. 2002. Genetic feature subset selection for gender classification: a comparison study. In *IEEE Workshop on Applications of Computer Vision*. IEEE Comput. Soc, 165–170.
[97] Amelia Thorpe. 2004. Where Do We Go? Gender Identity and Gendered Spaces in Postsecondary Institutions. *Antistasis* 7, 1 (2004), 1–12.
[98] Eileen M Trauth. 2013. The role of theory in gender and information systems research. *INFORG* 23, 4 (Oct. 2013), 277–293.
[99] Radu-Daniel Vatavu. 2017. Fundamentals of Gesture Production, Recognition, and Analysis. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1174–1177. https://doi.org/10.1145/3027063.3027106
[100] Susann Wagenknecht, Min Lee, Caitlin Lustig, Jacki O'Neill, and Himanshu Zade. 2016. Algorithms at Work: Empirical Diversity, Analytic Vocabularies, Design Implications. In *ACM Conference on Computer Supported Cooperative Work*. ACM Press, New York, New York, USA, 536–543.
[101] Candace West and Don H Zimmerman. 1987. Doing Gender. *Gender & Society* 1, 2 (June 1987), 125–151.
[102] Candace West and Don H Zimmerman. 2009. Accounting for Doing Gender. *Gender & Society* 23, 1 (Feb. 2009), 112–122.
[103] Baiqiang Xia, Boulbaba Ben Amor, Hassen Drira, Mohamed Daoudi, and Lahoucine Ballihi. 2015. Combining face averageness and symmetry for 3D-based gender classification. *Pattern Recognition* 48, 3 (2015), 746–758. https://doi.org/10.1016/j.patcog.2014.09.021
[104] Jian Yang, Lei Zhang, Yong Xu, and Jing Yu Yang. 2012. Beyond sparsity: The role of L1-optimizer in pattern classification. *Pattern Recognition* 45, 3 (March 2012), 1104–1118.
[105] Meng Yang, Lei Zhang, Xiangchu Feng, and David Zhang. 2014. Sparse Representation Based Fisher Discrimination Dictionary Learning for Image Classification. *International Journal of Computer Vision* 109, 3 (2014), 209–232. https://doi.org/10.1007/s11263-014-0722-8
[106] Ming Yang, Shenghuo Zhu, Fengjun Lv, and Kai Yu. 2011. Correspondence Driven Adaptation for Human Profile Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1, c (2011), 505–512. https://doi.org/10.1109/CVPR.2011.5995481
[107] Jang-Hee Yoo, Doosung Hwang, and Mark S Nixon. 2005. Gender classification in human gait using support vector machine. In *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 138–145.
[108] Emilio Zagheni, Venkata Rama Kiran Garimella, Ingmar Weber, and Bogdan State. 2014. Inferring International and Internal Migration Patterns from Twitter data. *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion* (2014), 439–444. https://doi.org/10.1145/2567948.2576930
[109] Yin Zheng, Richard S. Zemel, Yu Jin Zhang, and Hugo Larochelle. 2015. A Neural Autoregressive Approach to Attention-based Recognition. *International Journal of Computer Vision* 113, 1 (2015), 67–79. https://doi.org/10.1007/s11263-014-0765-x

Received April 2018; revised July 2018; accepted September 2018