



Queer perspectives could be used to gaze beyond the limiting binary and find ways to build AI technologies and systems for an inclusive future.



Beyond the Binary—Queering AI for an Inclusive Future

🔥 Evelina Liliequist, Andrea Aler Tubella, Karin Danielsson, and Copp  e Cocq, Ume   University

Nowadays it is somewhat outdated, perhaps even naive, to talk about artificial intelligence as something set in the future. AI systems are already integrated into everyday human life, although they're not the sci-fi-inspired robots some imagine as AI. In fact, there are AI systems that operate much closer to home, in or on the human body, to be more specific. A specific type of AI is facial analysis technology (FA), an application where a person's facial and bodily information such as face shape, features, skin color, movements, or makeup are captured, analyzed, and compared. This allows the development of systems that can match a recognized face against a database (face recognition), compare a recognized face against a given match (face authentication), or classify individuals (e.g., in terms of gender, race, scars, body geometry). Facial analysis technologies are already widely in use, for example, to allow you to unlock your phone just by scanning your face or to tag social media photos automatically. However, its use also includes more controversial applications, such as surveillance or recognition of crime suspects.

Many AI systems rely on labeling, that is, given an input, choosing from a selection of labels to describe it. This is also the case for FA tech, where the input is a face, and labeling is often based on binary systems, dividing data into categories such as man/woman, child/adult, and human/animal. In recent years, scholars in computer science, AI, and informatics have critically examined and highlighted problematic aspects of FA tech regarding race and gender [1,2,3,4]. Previous studies have identified key problems that raise concerns about how reliable or even how useful this technology can be. For example, the data used to train these systems is often not representative. In most cases, it is overwhelmingly white and male, which means that darker-skinned females have a disproportionate rate of misclassification, as has been found in comprehensive analyses of FA gender-classification systems [1]. Additionally, training datasets are often made up of images scraped from social media [5], which means (aside from privacy issues) that they capture only those who use social media, the specific

Insights

- Queer perspectives can help identify where risks concerning visibility and representation arise.
- Keeping a queer eye fixed on how AI systems are designed and trained, we aim at gazing beyond the limiting binary to find new ways of building AI technologies and systems for an inclusive future.

A queer perspective can help guide us into finding new ways for more-inclusive technology.

context of public photos, and whatever information people feel comfortable disclosing in their circles. But an image can represent many things—how we are expected to be viewed, how we see ourselves, and how we see ourselves and others beyond history’s assumptions, bias, and prejudices [6].

Often, the proposed solutions to these problems involve making datasets more representative (i.e., including more images of underrepresented groups) and/or involving at-risk groups in the design process. Although, of course, those affected by FA should have a say in how it is designed, these solutions themselves bring challenges. In particular, decisions about sexual orientation information, including what should be disclosed, where, and to whom can be made as an ongoing context-based process: This means that ripping images from their original context to use for training unrelated systems can violate their original purpose and intentions. Similarly, involving at-risk communities in any process poses a risk of exploiting their experience and requires disclosing information that they may not wish to share [5]. Thus, if technology is to allow for visibility and representation, we need to take the utmost care to ensure that this visibility and representation are not exploited and perverted into harmful systems. We argue that a queer perspective can help us identify where these risks arise, as well as guide us into finding new ways for more-inclusive technology.

DIVIDE AND CONQUER—THE ISSUE OF DATA LABELING

A key issue surrounding AI technologies such as FA is the labeling itself. In our forthcoming book chapter [7] and in our ongoing research, we argue that not all humans wish to identify themselves within a binary and/or perform their gender and/or sexuality in stable categories. Our arguments originate from queer theory (e.g., [8]): Labeling is an act of division, and as such, ultimately is often also an exercise of power. In a large body of scholarly work, queer theoretical thinkers have identified hetero- and cisnormative structures as deeply rooted in dichotomies, with intimate ties to power. Divisions into normal/deviant, good/bad, right/wrong, and so on are maintained and re-created by encouraging and rewarding the identities and practices that are categorized as normative and desirable, and by invisibilizing, restricting, and punishing other gender expressions, bodies, and sexualities [8]. To differ from the constructed “normal” can have serious consequences, such as being punished with external violence and/or becoming inflicted with internalized homo- and/or transphobia, which can lead to mental health issues and self-destructive behavior such as suicide.

Further, there is a great risk for not only reproducing, but also cementing hegemonic norms of gender identities and performances, and/or sexualities, based on stereotypes. In that sense, FA is a particularly egregious example: By its premise, it suggests that gender can be determined directly from what can be read from a face. This implication is based

on an essentialist understanding of sexuality as a fixed identity reflected in our body, without regards for self-identification. Tomasev et al. [9] argue that reinforcement of stereotypes can risk “echo[ing] tenets of eugenics—a historical framework that leveraged science and technology to justify individual and structural violence against people perceived as inferior.”

Against this background, where the labeling acts as a fundament for how AI technologies like FA work, there is a risk that already marginalized identities and bodies become even more vulnerable. For example, transgender individuals report an overwhelming critique and negative attitude toward automated gender recognition [9].

By finding ways of queering data, we can imagine and create a more inclusive future [2,6]. Grace Turtle argues, “to the binary logic of AI, differently situated knowledge(s) and other/ed perspectives, are destabilising, thus demonstrating that binaries, be they imagined, cultural or technological are not natural nor always necessary” [2]. Along similar lines, we argue that queer perspectives could be used to find new ways forward, and hopefully reach beyond data labeling within the limitations of binaries.

(A)I SPY WITH ITS LITTLE EYE—CONCERNS OF VISIBILITY AND RISKS FOR ANTI-LGBTQ USAGE

In the past two decades in a Swedish context, there have been many advances in queer socio-legal rights, which have led to a seemingly more open and queer-inclusive society. But looking back just a few decades, anonymity was an important and necessary part of the lesbian, gay, bisexual, trans, and queer community and queer lives, which can be understood in relation to the earlier, much harsher climate toward LGBTQ people. For example, until 1946, homosexuality was a criminal offense in Sweden according to chapter 18 of the Penal Code § 10, a law that forbade tidal law and “fornication against nature.” After decriminalization, homosexuality was still classified as a mental disorder by the World Health Organization until 1979. The National Board of Health and Welfare abolished the diagnosis “transvestism” as a psychiatric diagnosis as late as 2009, and it was not until 2013 that the requirement that transgender people be sterilized in order to receive gender-confirming care was deleted from the Gender Equality Act.

Despite an increase in queer visibility and acceptance in Swedish society, and several studies indicating an increased visibility of LGBTQ people in society, individuals may still feel a need to conceal their sexual orientation and/or gender identity or expression. One example is digital platforms, where LGBTQ people can feel a need to handle expectations of sharing and make strategic decisions around *what*, *where*, and *for whom* as disclosures regarding their sexual orientation. Sharing personal information about sexual orientation and expressions of being queer often heavily depends on the imagined audience on different social media sites. Such a strategy of selective self-presentation is akin to creating a “digital closet,” referring back to the concept of “the closet” as a metaphor for queer oppression, denial, and concealment. This also problematizes the binary logics of the closet—either you’re in or out—by instead pointing toward openness and visibility as an ongoing and context-based process [7].

Furthermore, the concept of selective self-presentation

immediately leads to questions on how images posted in—and meant for—specific contexts can be ripped from such contexts and used for other purposes. A highly debated study by Michal Kosinski and Yiland Wang [10] purports that classifiers can achieve a higher accuracy than humans in inferring sexual orientation from face images. This study used a training dataset of images obtained from online dating websites and validated on images obtained from Facebook, taken without explicit consent and validated for individuals who self-reported sexual orientation in a dating context or in a semipublic Facebook context. The possibility of data being used in such a way immediately leads to the question: What do we want to show online? If we dare to be visible in a certain way in a certain space, can that be taken from us and used to infer unrelated labels in a completely different context? The simple goal of having a system that identifies sexual orientation denotes a bigger problem in the machine learning space: the desire to draw strict categorizations in what is a fluid space.

Another issue that the critiqued study by Kosinski and Wang [10] raises is the question of usage. If such predictions of a person's sexuality actually *could* be made, for what would they be used? We argue that with such technology there is always a risk of anti-LGBTQ usages. Lest we forget, in large parts of the world, Sweden included, LGBTQ people are exposed to violence and discrimination. Further, homosexuality is still banned in many countries, in some under penalty of death. Even in cases where this type of prohibition legislation is rarely applied, it affects the social climate by legitimizing discrimination and violence against LGBTQ people, or people who are perceived to be LGBTQ.

QUEERING AI FOR AN INCLUSIVE FUTURE

Both because of its everyday applications and potential controversial uses, we argue that it is essential to analyze FA technology and its implications, with special care toward vulnerable groups in society. In 2022, Humlab (<https://showcase.humlab.umu.se/queer-ai>) at Umeå University organized a series of talks about queer perspectives on AI. The seminar discussions concluded that we, as designers, engineers, researchers, and citizens, need to understand the potential use and misuse of AI systems. Moreover, we must reflect on how we could and should design these systems, addressing possibilities, constraints, and risks from a variety of perspectives, including LGBTQ ones. Queer theory has great potential to assist us in addressing these issues. In our book chapter [7], we elaborate on this and argue for the necessity of casting a queer eye on AI to find new ways of relating to technology. In research we are currently planning, we will explore how AI systems can be designed to operate in a much more fluid and complex reality than what can be captured in a simple binary.

As a strong symbol of the future, AI is already integrated into society; the future is therefore already present. Still, every day brings new possibilities to change which directions and societal needs that AI technologies should be used for. Keeping a queer eye fixed on how AI systems are designed and trained, it is our most fervent hope to be able to gaze beyond the limiting binary and find new ways of building AI technologies and systems for an inclusive future.

ACKNOWLEDGMENTS

We would like to thank all the speakers, discussants, and audiences that took part in the Queer AI seminar series at Humlab, Umeå University, in 2022. The seminars were funded by the Faculty of Arts and Humanities, Umeå University. We would also like to highlight and celebrate the Queer in AI initiative that aims to raise awareness of queer issues in AI/ML, foster a community of queer researchers, and celebrate the work of queer scientists (<https://www.queerinaai.com>).

ENDNOTES

1. Buolamwini, J. and Gebru, T. Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proc. of the 1st Conference on Fairness, Accountability and Transparency*. PMLR 81 (2018), 77–91.
2. Turtle, G. Mutant in the mirror: Queer becoming with AI. *Proc. of DRS 2022*.
3. Rincón, C., Keyes, O., and Cath, C. Speaking from experience: Trans/non-binary requirements for voice-activated AI. *Proc. of the ACM Hum.-Comput. Interact.* 5, CSCW1 (Apr. 2021), Article 132, 1–27; <https://doi.org/10.1145/3449206>
4. Whittaker, M. et al. *Disability, Bias, and AI*. AI Now Institute, 2019.
5. Raji, I.D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., and Denton, E. Saving face: Investigating the ethical concerns of facial recognition auditing. *Proc. of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)*. ACM, New York, 2020, 145–151.
6. Matebeni, Z. Intimacy, queerness, race. *Cultural Studies* 27, 3 (2013), 404–417.
7. Danielsson, K., Tubella, A.A., Liliequist, E., and Cocq, C. Queer Eye on AI: binary systems versus fluid identities. In *Handbook of Critical Studies of Artificial Intelligence*. S. Lindgren, ed. Edward Elgar, Cheltenham, Forthcoming 2023.
8. Butler, J. *Gender Trouble: Feminism and the Subversion of Identity* (10th anniversary edition). Routledge, 1999.
9. Tomasev, N., McKee, K.R., Kay, J., and Mohamed, S. Fairness for unobserved characteristics: Insights from technological impacts on queer communities. *Proc. of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York, 2021, 254–265.
10. Kosinski, M. and Wang, Y. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology* 114, 2 (2018), 246.

👤 **Evelina Liliequist** is a postdoctoral researcher at the Centre for Regional Science at Umeå University. She is also an affiliated researcher at Humlab, Umeå University, and an affiliated researcher in the research cluster TechnAct, University of Gothenburg. Her research revolves around queer orientations, social media, space, and place.
→ evelina.liliequist@umu.se

👤 **Andrea Aler Tubella** is a senior research engineer in the Responsible AI group of Umeå University. Her research revolves around responsible AI, through the development of logical formalisms and tools, as well as education and sociotechnical aspects of AI systems.
→ andrea.aler@umu.se

👤 **Karin Danielsson** is an associate professor in the Department of Informatics at Umeå University. She is also director of Humlab, Umeå University. Her research revolves around participation during design and critical perspectives on design and evaluation of digital materials for different user groups in a variety of contexts.
→ karin.danielsson@umu.se

👤 **Coppélie Cocq** is a professor and deputy director of Humlab, Umeå University. Her research focuses on digital practices, critical studies in minority and Indigenous research, and ethical and methodological perspectives on digital research. She is an affiliated researcher in the research cluster TechnAct, University of Gothenburg.
→ coppelie.cocq@umu.se

