Patterns

CellPress
OPEN ACCESS

Opinion

Advancing LGBTQ+ inclusion in STEM education and AI research

Emily Wong,[1,2] Ryan J. Urbanowicz,[1,2] Tiffani J. Bright,[1,2] Nicholas P. Tatonetti,[1] Yi-Wen Hsiao,[1] Xiuzhen Huang,[1] Jason H. Moore,[1] and Pei-Chen Peng[1,*]
[1]Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[2]These authors contributed equally
*Correspondence: pei-chen.peng@cshs.org
https://doi.org/10.1016/j.patter.2024.101010

The authors emphasize diversity, equity, and inclusion in STEM education and artificial intelligence (AI) research, focusing on LGBTQ+ representation. They discuss the challenges faced by queer scientists, educational resources, the implementation of National AI Campus, and the notion of intersectionality. The authors hope to ensure supportive and respectful engagement across all communities.

Introduction

The history of science has highlighted the consequences of excluding minorities and women from clinical research. For example, up until 1994, women were not adequately represented in clinical trials for being perceived as more biologically complicated than men, which resulted in less generalizable knowledge and inadequate treatment options. Today, the misuse of sex and gender in health research leads to an incomplete scientific understanding of the relationship between one or both of those attributes and health outcomes (e.g., breast cancer) and thus insufficient medical care and health inequities.[1] While assuming a gender and sex binary has been a convenient methodological decision, it fails to capture the true complexity of gender and sex identities.

Statistics has been regarded as the grammar of all science, giving meaning to the underlying data. As machine learning (ML) and artificial intelligence (AI) technologies are increasingly being incorporated into clinical care and research, special attention should be given to addressing how such misrepresentations impact models and patients. Whether through traditional statistical modeling, ML, or AI, misrepresentation of data can yield misleading and potentially harmful conclusions.[2] At the root of the issue, the erasure of gender and sex diversity in STEM (science, technology, engineering, and mathematics) education inevitably bleeds into clinical research methods and thereby statistical, ML, and AI models. Moreover, LGBTQ+ (that is, lesbian, gay, bisexual, transgender, queer, and more) scientists are underrepresented in the biomedical AI community and STEM more broadly, which can stifle scientific innovation.[3] In addition to the known benefits of diverse groups for scientific problem solving, excluding scientists based on protected attributes such as sexual or gender identity can cost the scientific community great minds.

Throughout history, queer scientists, such as Alan Turing, have contributed monumental scientific innovations that have set the stage for modern developments. As the father of theoretical computer science and the inventor of the first modern computer, Alan Turing was prosecuted for his queer identity and later died of cyanide poisoning at just 41 years old. While there have been great improvements in encouraging greater inclusivity of LGBTQ+ in biomedical AI and STEM more broadly, many queer scientists continue to face career limitations, harassment, social exclusion, and discrimination.[3] Therefore, the remaining sections argue that the research community should (1) foster an environment that encourages and welcomes scientists from the LGBTQ+ community to engage with and lead biomedical and AI research and (2) encourage all scientists to take initiative in advancing our understanding of LGBTQ+ health issues.

Needs and strategies for LGBTQ+ inclusion in STEM education and AI research

Assessing needs and improving visibility

The LGBTQ+ community is often referred to as an invisible minority in STEM. A sparse collection of sexual orientation or gender identity information contributes to this invisibility as a result of historical bias against LGBTQ+ individuals and heteronormative stereotypes. In AI healthcare research, the lack of comprehensive surveys about gender and sexual identity makes it challenging to estimate the number of LGBTQ+ individuals and address healthcare inequities. Collecting and analyzing data that support cultural and identity information while maintaining scientific rigor is essential. In addition to data collection, we also recommend assessing the performance of AI models to ensure health equity for LGBTQ+ individuals. For example, the health equity assessment for machine learning framework[4] ensures AI tools are fair in health outcomes. It focuses on identifying factors linked to health inequities and setting performance metrics directly related to health equity to evaluate the fairness of AI models.

To establish inclusive environments in biomedical research and increase LGBTQ+ representation in STEM, we must first understand the current representation of queer individuals in the field by enhancing data collection practices. We strongly encourage clinicians and researchers to identify the specific needs of the LGBTQ+ community and advocate for their inclusion in healthcare and AI research.

Policies

We recommend reviewing policies to ensure inclusivity. For example, implementing pronoun inclusion policies can help remove communication barriers, recognize diverse gender identities, and promote inclusivity. The Trevor Project National Survey (https://www.thetrevor

project.org/survey-2024/) has shown that about 66% of LGBTQ+ youth use pronouns outside of binary options. Moreover, many STEM organizations, such as the European Molecular Biology Laboratory,[5] have started to increase the representation of LGBTQ+ individuals in leadership positions, advisory boards, and speaker panels. These initiatives not only encourage LGBTQ+ individuals to lead research efforts but also provide individuals with visible pathways to success in STEM careers. Policy revision can help foster inclusion and address disparities faced by LGBTQ+ individuals in STEM.

LGBTQ+ inclusion in STEM education

With proper education and training that encourages awareness of allyship practices, scientists will be better equipped to support their colleagues with diverse sexual orientation, gender identities, and expressions. Through a systematic literature review and inductive thematic synthesis, Wright and Delgado[6] proposed a framework for creating gender and sexual diversity-inclusive STEM curricula. Particularly, this framework would encourage educators to do the following:

(1) Challenge heteronormativity. Popular biology textbooks often frame sexuality in terms of reproductive success, conflate gender and sex, and imply a sex and gender binary model. Educators can encourage students to interrogate assumptions of a gender binary and opt to use gender-neutral language in the classroom.
(2) Center the curriculum on advancing social justice, which requires legitimizing the experiences of queer people. Educators can support transgender students in biology by incorporating a diversity statement affirming the faculty member's support or asking students to provide their preferred pronouns.
(3) Reaffirm the complexity of identities. Specifically, that identity is a fluid, intersectional construct influenced by both biological and nonbiological factors.

Acknowledging intersectional identities

Intersectionality is crucial for fostering inclusive environments. Originally described by Crenshaw,[7] it recognizes that Black women, for example, face specific forms of discrimination that Black men or White women might not. Applying an intersectional lens fosters inclusive environments by understanding these varied experiences and identifying opportunities for improvement. This is particularly crucial for LGBTQ+ individuals, as it acknowledges how race, gender, sexual orientation, disability, and other factors intersect to shape unique experiences of discrimination and privilege.[8] In addition to the recommendations put forth by Wright and Delgado,[6] embracing intersectionality in biomedical and AI research involves establishing inclusive AI training programs that support LGBTQ+ participation and retention.

LGBTQ+ inclusion in AI research

AI now pervades numerous aspects of everyday life, including healthcare and bioinformatics. As such, every individual and community will be impacted. However, the impacts of AI are not uniform across communities. AI models learn from massive amounts of training data that often have encoded in them systematic biases and social norms. These biases and norms are learned and perpetuated by AI, increasing the risk of harm for marginalized groups.[9] Specific to the LGBTQ+ community, cis- and heteronormativity ("cishet") are deeply ingrained in natural language and other data used to train models. In healthcare, LGBTQ+ individuals are underrepresented or mislabeled in health data, which can negatively impact patient outcomes. Techniques to improve model performance for queer individuals can include prompt engineering, fine-tuning, or diversity-informed model development (e.g., deciding how to encode gender in prediction models). However, queer issues are grossly ignored in the AI domain and STEM more broadly. To increase and maintain LGBTQ+ representation in AI, we suggest inclusive AI education as a medium to integrate queer perspectives and active participation in the model development process as well as encouragement for all AI scientists to learn about queer issues to move away from the "cishet" baseline.

Existing resources for LGBTQ+ inclusion in STEM education and AI research

The availability and visibility of LGBTQ+ STEM educational and professional resources can serve to shape the attitudes and behaviors of both LGBTQ+ and the general scientific community. This includes promoting inclusivity and diversity in field membership as well as promoting awareness of unique LGBTQ+ challenges impacting AI research. Further, integrating intersectional approaches into AI education is crucial.

There are multiple educational resources available to help build a more inclusive community in STEM. Conferences are great venues for professionals to connect and discuss new policies. For example, the Grace Hopper Celebration is the world's largest gathering for women and non-binary technologists, offering a combination of technical sessions and career sessions. Unlike traditional research-focused conferences, the Grace Hopper Celebration offers round-table discussions and networking events tailored for LGBTQ+ professionals and allies. Similarly, the Annual Biomedical Research Conference for Minoritized Scientists is a premier conference for historically excluded communities in STEM. It provides a platform for scientists to share their research achievements at all career stages. The Tapia Conference promotes and celebrates diversity in computing, bringing together individuals from diverse backgrounds, including those with disabilities or of different genders, ethnicities, and races. All these conferences provide travel awards or scholarships to LGBTQ+ individuals and historically underrepresented trainees.

Professional associations also offer educational resources and networking opportunities for LGBTQ+ individuals in AI. Queer in AI (https://www.queerinai.com/), for instance, aims to raise awareness of queer issues in AI. The organization provides essential resources, including a dedicated Queer in AI Safety Team to address any member's experience of harassment or hostile behavior, as well as guidelines to ensure virtual conferences are queer friendly. Queer in AI also hosts social events, advocacy efforts, and workshops in major computer science conferences such as the International Conference on Machine Learning and the North American Chapter of the Association for Computational Linguistics. Other organizations, such as Out in STEM (https://www.ostem.org/) and Out to Innovate (https://noglstp.org/), formerly

known as the National Organization of Gay and Lesbian Scientists and Technical Professionals, focus on fostering inclusivity in STEM. Out in STEM empowers LGBTQ+ individuals in STEM through networking opportunities, mentorship, and resources. Out to Innovate provides professional development, networking, and peer support to LGBTQ+ individuals in STEM and allies.

Inclusivity at the National AI Campus

The National AI Campus is a project-based training program aiming to make AI accessible to a broad community. Since its establishment in 2018, the program has fostered learning collaboration between faculty, physicians, postdocs, research staff, and students from high school to graduate levels. The program is free and provides an introductory experience in AI/ML to people from diverse educational, ethnic, gender, racial, and socioeconomic backgrounds. The National AI Campus currently collaborates with 21 historically black colleges and universities and diverse academic programs across the US.

At the Cedars-Sinai Medical Center, we have adapted the national program to emphasize a greater diversity of participant degrees and training experience, with a stronger focus on biomedical projects (https://cedars.nationalcampus.ai/). For example, we offer projects that use AI to predict disease outcomes or develop a more reliable and efficient screening method for bladder cancer. The Cedars-Sinai AI Campus warmly welcomes participants from any academic background or specialty and any level of experience. By implementing an inclusive AI training program, the Cedars-Sinai AI Campus fosters a highly interactive environment where everyone can learn from each other. Beyond inclusivity, this program could be further extended to bring practical awareness to LGBTQ+-related issues and challenges by developing or inviting proposed AI/ML projects that specifically tackle LGBTQ+ health and/or analytical challenges.

Summary: Challenges and opportunities

Clearly, challenges remain for LGBTQ+ inclusion in STEM education and AI research. For example, the lack of a standardized LGBTQ+ curriculum hinders efforts to integrate respective topics into coursework. Additionally, there may be resistance or lack of awareness among educators and institutions about the importance of LGBTQ+ inclusion. Limited funding or incentives for LGBTQ+ initiatives can also make it difficult to develop and maintain comprehensive educational programs.

However, there are opportunities for collaboration and partnership to help overcome these challenges. Academic institutions can collaborate with LGBTQ+ advocacy organizations to develop inclusive curricula and training materials. Industry partnerships can provide real-world applications and case studies that demonstrate the relevance of LGBTQ+ inclusion in AI. Moreover, fostering allyship, support networks, and LGBTQ+ role models and mentors within STEM fields is crucial. By promoting inclusive environments and advocating for LGBTQ+ representation, allies can help create a more welcoming and supportive community for LGBTQ+ individuals in STEM, leading to a more diverse and inclusive AI workforce and, subsequently, more innovative and equitable technological advancements.

DECLARATION OF INTERESTS

Jason H. Moore serves on the advisory board of *Patterns*.

REFERENCES

1. Morrison, T., Dinno, A., and Salmon, T. (2021). The erasure of intersex, transgender, nonbinary, and agender experiences through misuse of sex and gender in health research. Am. J. Epidemiol. *190*, 2712–2717. https://doi.org/10.1093/aje/kwab221.
2. Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S. (2019). Dissecting Racial Bias in an Algorithm used to Manage the Health of Populations. Science *366*, 447–453. https://doi.org/10.1126/science.aax2342.
3. Cech, E.A., and Waidzunas, T.J. (2021). Systemic inequalities for LGBTQ professionals in STEM. Sci. Adv. *7*, eabe0933. https://doi.org/10.1126/sciadv.abe0933.
4. Schaekermann, M., Spitz, T., Pyles, M., Cole-Lewis, H., Wulczyn, E., Pfohl, S.R., Martin, D., Jaroensri, R., Keeling, G., and Liu, Y. (2024). Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study. EClinicalMedicine *70*, 102479. https://doi.org/10.1016/j.eclinm.2024.102479.
5. Unsay, J.D. (2020). LGBTQ+ in STEM: Visibility and Beyond. Chem. Eur. J. *26*, 8670–8675. https://doi.org/10.1002/chem.202002474.
6. Wright, G.W., and Delgado, C. (2023). Generating a framework for gender and sexual diversity-inclusive STEM education. Sci. Educ. *107*, 713–740. https://doi.org/10.1002/sce.21786.
7. Crenshaw, K. (1991). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics [1989]. In Feminist Legal Theory: Readings in Law and Gender, K.T. Bartlett and R. Kennedy, eds. (Routledge), pp. 57–80. https://doi.org/10.4324/9780429500480-5.
8. Buolamwini, J., and Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In Proceedings of the 1st Conference on Fairness, Accountability and Transparency, *81*, S.A. Friedler and C. Wilson, eds. (PMLR), pp. 77–91.
9. McAra-Hunter, D. (2024). How AI hype impacts the LGBTQ + community. AI Ethics. https://doi.org/10.1007/s43681-024-00423-8.

About the authors

Emily Wong: Dr. Emily Wong is a research data scientist in the Department of Computational Biomedicine at the Cedars-Sinai Medical Center and works closely with Dr. Tiffani Bright to develop responsible AI for healthcare. Her research focuses on integrating known human cognitive constraints and moral values into machine learning and generative AI systems.

Ryan J. Urbanowicz: Dr. Ryan Urbanowicz is an assistant professor of computational biomedicine at the Cedars-Sinai Medical Center, an adjunct assistant professor of informatics at the University of Pennsylvania, and director of the Cedars AI-Campus Program. His research focuses on the development, evaluation, and application of machine learning and artificial intelligence methods for the analysis of biomedical and clinical data. This includes specialization in interpretable modeling with rule-based machine learning, the adoption of best practices in data science within automated machine learning frameworks, and evolutionary algorithms for data driven complex associations discovery. He is a dedicated educator and mentor, making new educational materials available on his lab's YouTube channel (The URBS-Lab).

Tiffani J. Bright: Dr. Tiffani J. Bright is an assistant professor at the Cedars-Sinai Medical Center in the Department of Computational Biomedicine and co-director of the Center for Artificial Intelligence Research and Education. Dr. Bright is a distinguished expert in applied clinical informatics, particularly in the design, evaluation, and optimization of clinical decision support systems, and a recognized authority in AI ethics. She has championed diversity and inclusivity in informatics through both local and national initiatives and is an elected fellow of the American College of Medical Informatics.

Nicholas P. Tatonetti: Dr. Nicholas Tatonetti is a professor and vice chair of the Department of Computational Biomedicine at the Cedars-Sinai Medical Center and associate director of computational oncology at Cedars-Sinai Cancer. Dr. Tatonetti's research focuses on the discovery and validation of unknown and unexpected adverse drug reactions and drug-drug interactions using large-scale analyses and real-world data. They use AI/ML methods to advance our understanding of pharmacological

safety, especially for historically underserved and understudied populations in medicine. Recent work focuses on the use of foundation models for the integration of language, biomedical imaging, and multi-omics for the study of adverse event mechanisms of action, and they are an elected fellow of the American College of Medical Informatics.

Yi-Wen Hsiao: Dr. Yi-Wen Hsiao is a postdoctoral scientist in the Department of Computational Biomedicine at the Cedars-Sinai Medical Center. His primary research focus lies in using bioinformatics and biostatistics techniques to analyze genomic and healthcare data, particularly in exploring the complexities of human cancer. Currently, he is also working on evaluating health equity within these data, with the goal of addressing disparities in medical fields.

Xiuzhen Huang: Dr. Xiuzhen Huang is currently a professor in the Department of Computational Biomedicine at the Cedars-Sinai Medical Center. Prior to joining Cedars-Sinai Medical Center in fall 2022, Dr. Huang was a professor of computer science at Arkansas State University for 18 years. Her research interests include bioinformatics and biomedical informatics; medical imaging and genomics; artificial intelligence, machine learning, and deep learning; graph theory and algorithms; parameterized computation and complexity; and theory of computation.

Jason H. Moore: Dr. Jason Moore is chair of the Department of Computational Biomedicine at the Cedars-Sinai Medical Center, where he serves as director of the Center for Artificial Intelligence Research and Education. He leads an active research program focused on the development and application of AI algorithms for the analysis of biomedical data. Recent work focuses on methods for automated machine learning (AutoML) with the goal of democratizing AI in healthcare. He is an elected fellow of the American College of Medical Informatics, the International Academy of Health Sciences Informatics, and the American Statistical Association.

Pei-Chen Peng: Dr. Pei-Chen Peng is an assistant professor in the Department of Computational Biomedicine at the Cedars-Sinai Medical Center. Her research focuses on human genetics, cancer genomics, and statistical modeling. Dr. Peng also serves on the Equity and Inclusion Committee of the Genetics Society of America, recommending policies and practices to advance program goals. Throughout her career, she has been dedicated to teaching and mentoring computational biologists who enhance our understanding of life and health.