Computational audits combat disparities in recognition

Authors

- **John Doe**
  [XXXX-XXXX-XXXX-XXXX](#) · [johndoe](#) · [johndoe](#)
  Department of Something, University of Whatever · Funded by Grant XXXXXXXX

- **Jane Roe**
  [XXXX-XXXX-XXXX-XXXX](#) · [janeroe](#)
  Department of Something, University of Whatever; Department of Whatever, University of Something

Recognition can greatly affect the trajectory of a scientist's career since each form of recognition paves the way to future accolades. Regardless of an individual's merit, biases can skew which scientists are recognized. Audits can combat this by identifying overlooked groups and holding the recognizing body accountable. This is exemplified in a recent study of disparities in scientific journalism, which found that predicted East Asian names were under-quoted and under-mentioned [1].

Auditing efforts can reduce disparities. Journalists found that self-audits of gender-based representation in their articles can help them approach parity [2/,3/]. In award recognition, an audit of the International Society for Computational Biology (ISCB) honorees revealed significant under-representation of people with predicted East Asian name origins and over-representation of US-affiliated scientists. After this study was publicized, ISCB's following honorees had the highest mean predicted probability of having an East Asian name of any previous year, and the nominating committee inducted their first China-based ISCB Fellow [4].

Ideally, audits would only use self-reported gender, ethnicity, and other identifications [5]. While this is possible prospectively, surveying is impractical for large groups and often impossible retrospectively [6/]. In contrast, computationally-derived predictions allow for large-scale audits with as broad a scope as needed. Numerous tools exist to algorithmically infer gender, nationality, and ethnicity using the feature most likely to be present in datasets: an individual's name [7,8/].

Prediction models are not a panacea; several factors limit both their accuracy and utility. For instance, gender associations of a given name can vary by culture, potentially biasing gender predictions where additional information is not available [9]. Also, most gender prediction models are trained on binary gender labels, which systematically overlooks and conceptually erases the representation and contributions of transgender, non-binary, and intersex individuals [10].

Proxy predictions of ethnicity via name origin are more difficult still; choosing categories to probabilistically predict on is non-trivial and difficult to discretize. Furthermore, there is no one-to-one mapping between having a name from a linguistic group and belonging to a minoritized or underrepresented group. Colonialism, immigration, and structural racism have affected linguistic histories and inclusion or exclusion from scientific communities in complex ways that are nearly impossible to parse from names alone. For instance, classifiers are usually unable to distinguish if names of Hispanic origin come from the Iberian Peninsula or from Latin America [4].

Recognizing the aforementioned shortcomings, we propose recommendations for the creation and deployment of automated auditing tools: 1. *Transparency.* Publicly provide all tools, code, and data used in the analysis. This enables public scrutiny to those being audited and those whose data are used. For confidential data, provide de-identified or aggregated data.

2. *Individuals know best.* Self-identified demographic information should always be used in preference to algorithmic predictions.

3. *Aggregates only.* Audit results should not affect individuals. Analyses must focus on aggregate estimations and any intermediary individual predictions should not be used externally. In addition, analysts must be mindful of hidden subpopulations that may be obscured during aggregation.

4. *Inform the public.* While internal audits can help institutions reflect, researchers should inform the community of identified disparities. This should include a discussion of any caveats associated with the analyses performed (e.g., a binarization of gender) and the ethical implications of such choices.

References

1. **Analysis of scientific journalism in <i>Nature</i> reveals gender and regional disparities in coverage**
   Natalie R Davidson, Casey S Greene
   Cold Spring Harbor Laboratory (2021-06-22) https://doi.org/gkscd5
   DOI: 10.1101/2021.06.21.449261

2. **I Analyzed a Year of My Reporting for Gender Bias (Again)**
   Adrienne LaFrance
   The Atlantic (2016-02-17) https://www.theatlantic.com/technology/archive/2016/02/gender-diversity-journalism/463023/

3. **I Spent Two Years Trying to Fix the Gender Imbalance in My Stories**
   Ed Yong
   The Atlantic (2018-02-06) https://www.theatlantic.com/science/archive/2018/02/i-spent-two-years-trying-to-fix-the-gender-imbalance-in-my-stories/552404/

4. **Analysis of scientific society honors reveals disparities**
   Trang T Le, Daniel S Himmelstein, Ariel A Hippen, Matthew R Gazzara, Casey S Greene
   Cell Systems (2021-09) https://doi.org/gmhq49
   DOI: 10.1016/j.cels.2021.07.007 · PMID: 34555325

5. **Ten Simple Rules to Achieve Conference Speaker Gender Balance**
   Jennifer L Martin
   PLoS Computational Biology (2014-11-20) https://doi.org/gf853n
   DOI: 10.1371/journal.pcbi.1003903 · PMID: 25411977 · PMCID: PMC4238945

6. **How to Ethically and Responsibly Identify Gender in Large Datasets**
   MediaShift
   (2014-11-21) http://mediashift.org/2014/11/how-to-ethically-and-responsibly-identify-gender-in-large-datasets/

7. **Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records**
   Kosuke Imai, Kabir Khanna
   Political Analysis (2017-01-04) https://doi.org/f8ntmv
   DOI: 10.1093/pan/mpw001

8. **Genderize.io | Determine the gender of a name** https://genderize.io/

9. **A Data-driven Approach to Studying Given Names and their Gender and Ethnicity Associations**
   Shervin Malmasi
   Proceedings of the Australasian Language Technology Association Workshop 2014 (2014-11)
   https://aclanthology.org/U14-1021

10. **Racial and ethnic imbalance in neuroscience reference lists and intersections with gender**
    Maxwell A Bertolero, Jordan D Dworkin, Sophia U David, Claudia López Lloreda, Pragya Srivastava, Jennifer Stiso, Dale Zhou, Kafui Dzirasa, Damien A Fair, Antonia N Kaczkurkin, …
    Danielle S Bassett
    Cold Spring Harbor Laboratory (2020-10-12) https://doi.org/gj7mdc
    DOI: 10.1101/2020.10.12.336230