DV3-Safety 1.0 Perspective API (Google) **HateBERT** DV3 8.0 **Toxicity Probability** 0.6 0.4 0.2 0.0 Physical Disability Mental Disability LGBTQ Native American Middle East Women [†] Black [⊥] Muslim [→] Asian -Mexican Jewish *Chinese* [⊣] Latino