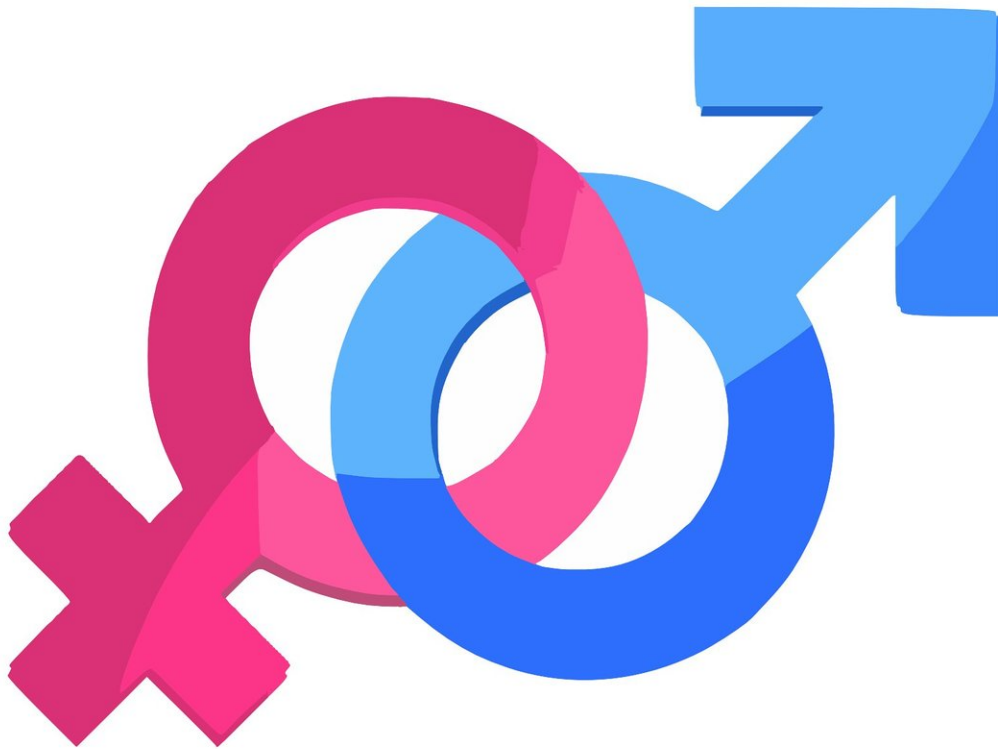# Examining gender stereotypes embedded in natural language

November 22 2023

Credit: CC0 Public Domain

Gender stereotypes harm people of both genders—and society more broadly—by steering and sometimes limiting people to behaviors, roles, and activities linked with their gender. Widely shared stereotypes include the assumption that men are more central to professional life while women are more central to domestic life. Other stereotypes link men with math and science and women with arts and liberal arts.

Perhaps surprisingly, research has shown that countries with higher economic development, individualism, and gender equality tend to also have more pronounced gender differences in several domains, a phenomenon known as the gender equality paradox.

To help explain this pattern, Clotilde Napp used a natural language processing model to look for stereotypes in large text corpora from more than 70 countries. The paper is published in the journal *PNAS Nexus*.

Napp's model looked for words representing the categories men and women as well as sets of words representing the attributes career-family, math-liberal arts, and science-arts. The model then applied the Word Embedding Association Test (WEAT), which measures the association between sets of target words in terms of their relative semantic similarity to sets of attribute words.

Napp finds that gender biases about careers, math, and science are all stronger in the text corpora of more economically developed and individualistic countries.

The author urges caution in interpreting the results which are based on big data analysis in an international context and may involve various underlying mechanisms. The cause of this pattern remains to be established with certainty, but Napp points to theoretical work suggesting that in societies where beliefs about the inherent inequality of men and women have declined, beliefs about the equality but inherent

differences of men and women may have emerged to replace older hierarchical ideas.

Another explanation, which is not mutually exclusive with the previous explanation, is that the biased associations reflect existing gender differences in behaviors that are stronger in wealthy countries.

The presence of gender stereotypes in the online text corpora used to train AI could reinforce these stereotypes in artificial intelligence models, according to the author.

More information: Clotilde Napp et al, Gender stereotypes embedded in natural language are stronger in more economically developed and individualistic countries, *PNAS Nexus* (2023). DOI: 10.1093/pnasnexus/pgad355

Provided by PNAS Nexus