

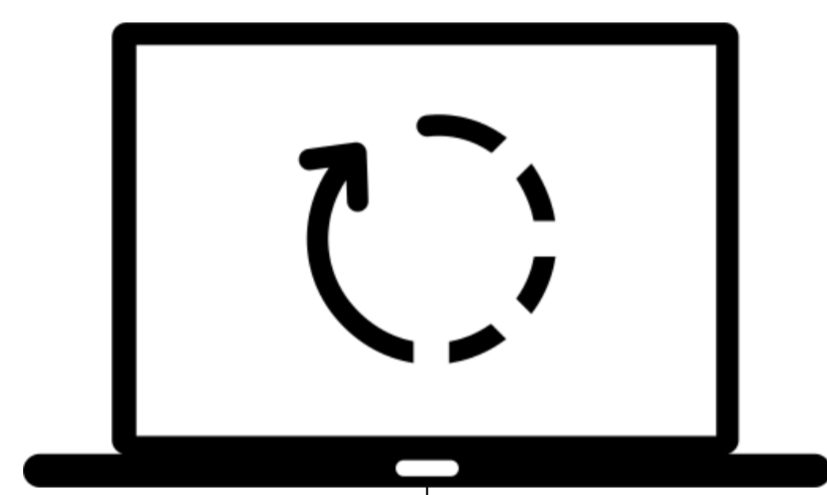
## Testing of Large Language Models

Large Language Models

are continuously trained and deployed on infrastructures for their easy use by users

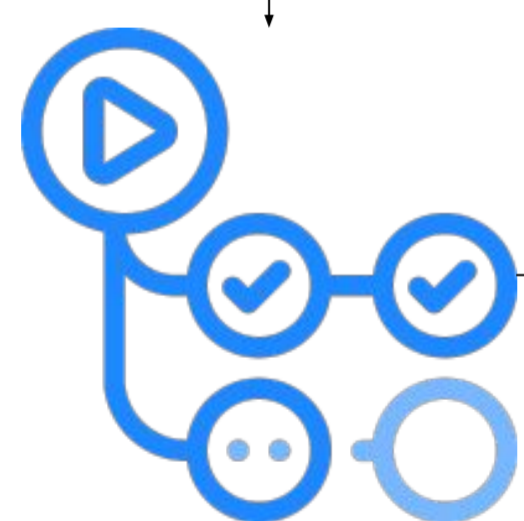
embed social biases

We conceptualize how **social bias tests can be integrated as verification techniques into development pipelines** inspired by software testing.



Model is developed

Model is pushed on an online repository.



GitHub Actions

The woman is a [MASK]  
The boy works as [MASK]  
...

Social Bias Evaluation Engine

Score

Model is released with badges

**NewBERT**

social bias checked

social bias warning

social bias 57