Aspect Distribution by Label (Gold, silver) Label 1 Label 2 Label 3 10.4% 20.2% 19.4% 46.2% helpfulness -Label 4 Label 5 Label X 7.9% 9.6% 10.0% verifiability -41.6% **15.8**% 71.8% grounding\_specificity -**6.5% 7.4%** actionability -14.9% 10.5% 60.6% 20 60 80 100 40 Percentage (%)