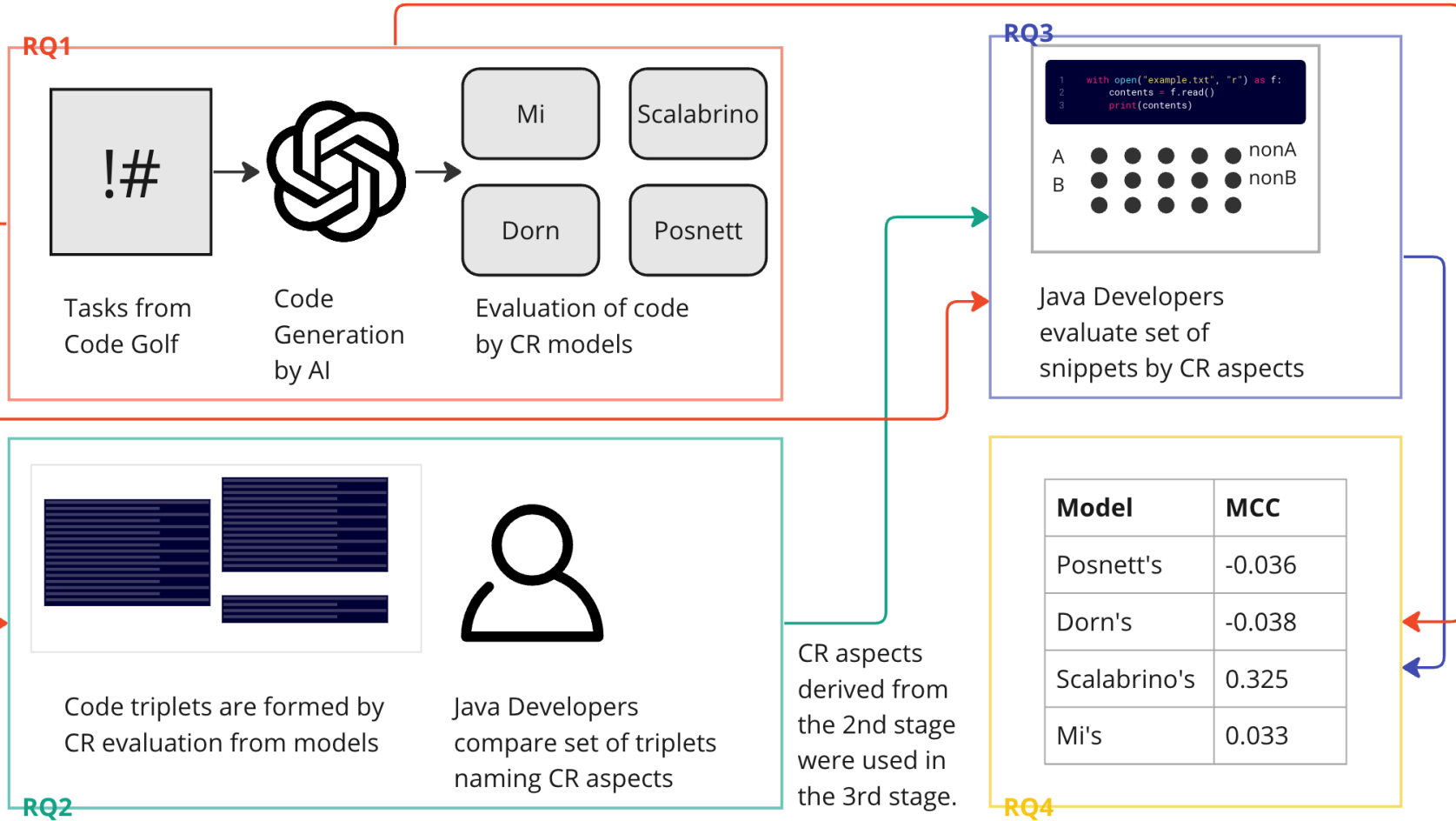


Snippets from the 1st stage were used in the 2nd and 3rd stages.



Evaluations from models from the 1st stage and human evaluations from the 3rd stage were used in the 4th stage.

CR aspects derived from the 2nd stage were used in the 3rd stage.