user

2 User specifies AI behaviors

1 User describes AI deployment scenario

6 User reviews (and optionally summarizes) examples of harm

Ethical Matrix

AI behaviors / stakeholders	false positive	false negative	...
the applicant			
the company			
the hiring manager			
....			

Vignette

Scenario

A company wants to deploy an AI hiring system to scan the resumes of applicants and predict whether they are a good fit for a given job opening.

Imagine you are an applicant for a job opening at the company. If the system determines that you are not qualified for a position when you are, you may be harmed because ...

Stakeholder — an applicant

AI behavior — you are not qualified for a position when you are

Generated harm example

... you miss out on the benefits and career opportunities associated with it.

Generated harm example

... I would have to look for another job and spend more time and resources for that. I would also start to doubt if I'm actually qualified for the job and similar ones to it. Depending on how the search goes, I'd either give up or just switch to a different field. The latter would require even more time and resources on my part.

Harm categories

▸ Allocational
▸ Legal & reputational
▸ Loss of rights & agency
▸ Quality of service
▸ Social & societal
▸ Representational
▸ Well-being
▸ Other

3 **LLM** AHA! generates stakeholders based on the scenario description

4 **script** AHA! generates vignettes based on the scenario, stakeholder, and AI behavior

5 **crowd** **LLM** AHA! generates examples of harm