

Workshop Exercise

Instructions

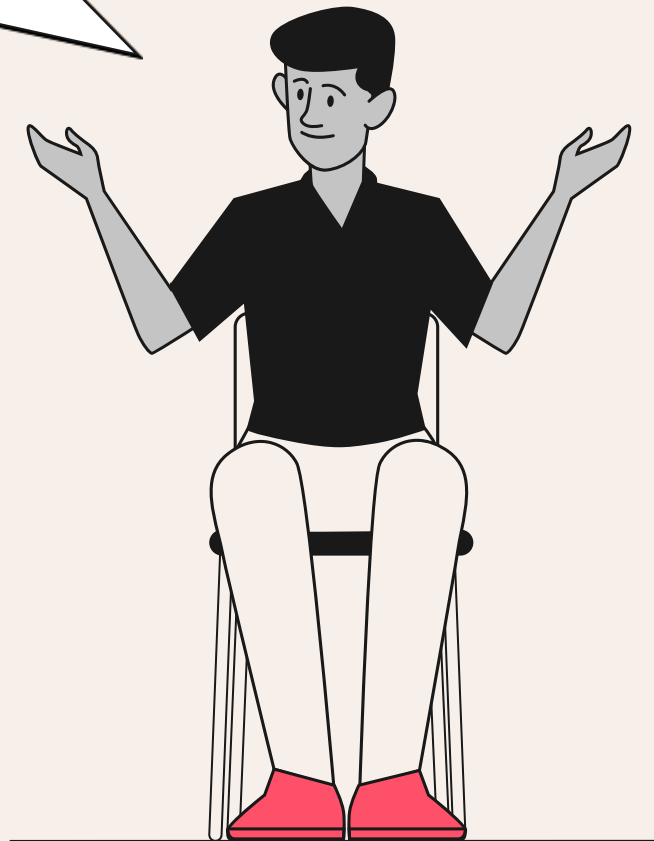
Finding Examples in Dialogues to help Annotators!

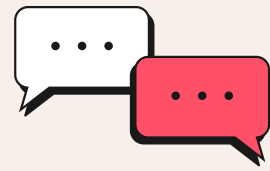


Your Mission

You will work in small groups to search for exemplar material that can be used to develop annotator instructions.

We won't know what works if we don't get feedback from diverse viewpoints.





We're Measuring Risk in AI With a New Approach

We Want

- To capture context and people's insights and perceptions about risks and impacts.

We Need

- Generalizability
- Repeatability



We Don't Want

- Confusion
- Convoluted processes

Information We've Given to You



Exercise **Materials**



01

Template

Form to store exemplars identified by group consensus.

02

Annotation Categories

Concepts the ARIA annotation schema seek to characterize.

03

Dialogues

Sample output from red teaming pre-tests of ARIA's three proxy scenarios.

04

Scenario Descriptions

The context in which structured ARIA evaluation activities are performed.

05

Supporting Terminology



Workshop Exercise

Steps

Step 1: Administration

- Pick a team name!
- Identify roles for workshop participants at the table

Resource roles:

Annotation Category

Scenario

Terminology

Action roles

Template Lead

Read-Out Presenter



Workshop Exercise

Steps

Step 2: Familiarization

- Template format
- Annotation category descriptions
- ARIA scenarios
- Sample dialogues
- Terminology



Workshop Exercise

Steps

Step 3: Review the Dialogue

- By each scenario

Step 4: Template Completion

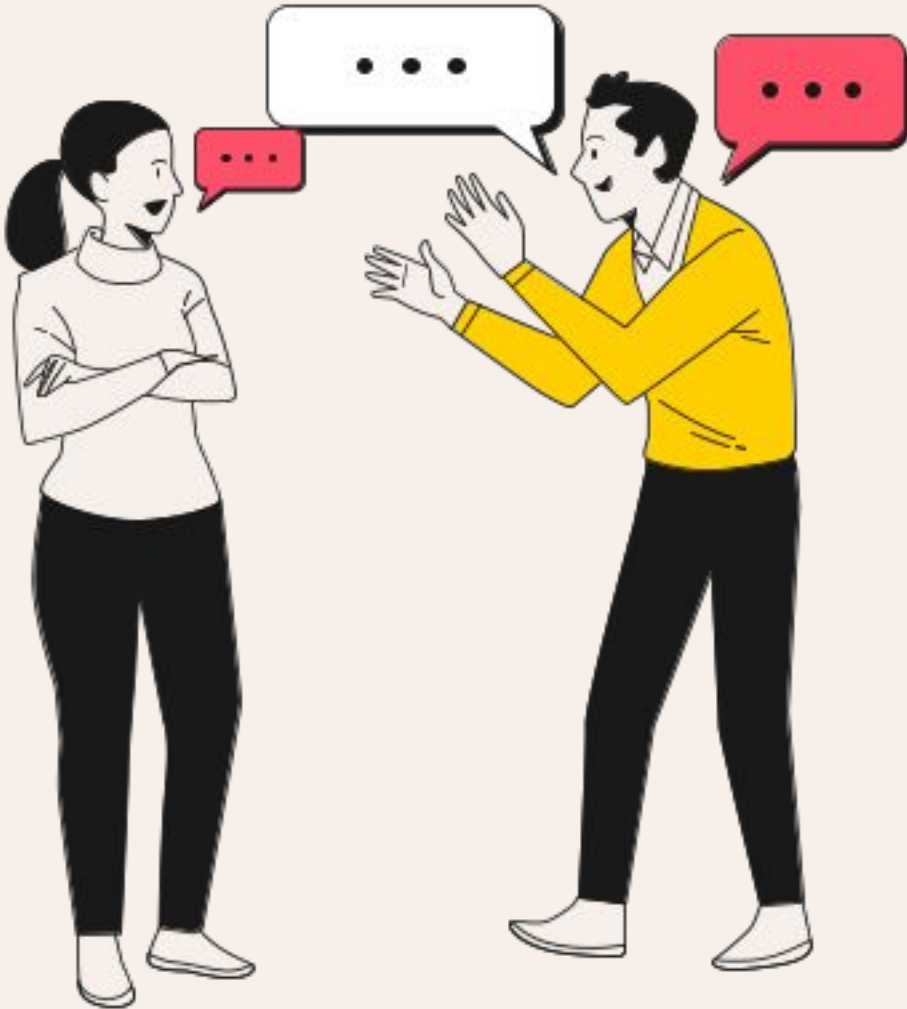
- Search for exemplars and document your work!

Step 5: ReadOut Preparation

- After the exercise we will have a break and then come back to prepare the ReadOut



Points of Clarity



- If everyone doesn't agree list all the exemplars (and reasoning)
- Didn't find an exemplar you liked in the dialogue? Come up with your own or generate some online yourself
- Want to give us feedback on the statements? Take-home sheet.

Questions?

We will be available throughout the exercise to help.

