

## **Expert assessment informal analysis report (Contestable Camera Cars)**

Aimed at identifying points of improvement for the video.

### **Overall summary**

Based on participant summaries below.

1. Most struggled to recognize value change element. However, adding this would require significant rework. The narrative related to contestation can also be told without introducing the issue of value change.
2. Many asked about some sort of voice-over. Time and budget constraints prevent us from adding one in. We will have to rely on more use of on-screen text and graphics to make the narrative as obvious as possible.
3. We need to more clearly connect the risks shown with the contestation procedure, and from there with the system improvements.
4. We need to show actual disagreement between people, mainly in the contestation procedure. Make it less polite, and more emotional. When a decision is made, not everyone has to be happy.
5. Show how AI is part of the process of contestation (e.g. in the scene with the debate, also show digital tools that both sides use to investigate AI behavior).
6. Show how contestation feeds back into ongoing system design & development (e.g. show a bunch of engineers in a room receive a message with the outcome of the appeals procedure and get to work).
7. Show that it's not (only) about individual grievances but also communities organizing around shared issues. Maybe at least in one case of contestation show a group rather than a single person talking to the city, via a spokesperson.

### **Participant summaries**

Only focused on the things participants found lacking or unclear.

#### **P1**

- Sees values but does not see them change

- Appeal procedure is abstract. Not clear what decision needs appealing in the first place. Did not see someone directly appealing an AI decision.
- Adversarial: did not see a *dispute*, but did see a discussion
- Human-in-the-loop: not clear. Thinks it's a person at home.
- Chilling effect: thinks it's people running away because they are afraid of getting caught
- Missed the model bias issue entirely
- Contestability loop is abstract, who is contesting what, and why?
- We should try to connect the people from the issues with the contestation process

## P2

- Dislikes that the garbage worker is a black person – avoid stereotypes
- Social values did not come up. The issues are functional. It shows the tech can be problematic, but not in terms of values. The video does not ask about values in terms of preferences.
- We should show more frustration around the appeal procedure
- Adversarial: show people object to the issues shown, and show how people disagree
- Contestation process shown is not speculative. Does not open up new ways of seeing. This is not strange familiarity. It is too familiar, not strange enough
- Show problems and contestations in succession. Allows for showing more variety of contestation.
- Values complicates things and makes it abstract. Need to invest much more in dialogue of characters to show preferences and beliefs. And ask questions about if things are fair and right.

## P3

- Not sure IA bias is opposed to a societal value?
- Feels the video does not show changing values
- Saw an appeal but not a *procedure*, did not recognize a step-based process
- The process wasn't adversarial. It is not clear what they were talking about, and if it was a dispute.
- There should be tension and adversity
- What's different from the current practice of contestation?
- We should improve clarity, by adding text or voice
- "Adversarial" makes one think of emotional aspects, whereas it's about a conversation
- "Dialectical" (a term we shared as synonymous with adversarial) *is* a

conversation (but also involves opposing forces)

- Of interest are at-scale effects, not just individual grievances
- Emphasize in-band nature of contestability we are proposing

#### **P4**

- Needs more support with text or voice-over
- Does not see value change
- Is this an example that really shows something where things are at stake?
- Value change would require showing how people's views of what trash is changes, and that the AI needs to adapt to that

#### **P5**

- Appeal procedure went by quickly
- Not clear when the appeal is happening in the larger process
- Interprets the human in the loop as part of appeal, not clear who it is
- Not clear what they shake hands on
- Collective level isn't clear from the video, only the individual
- Where in the process shown does the handshake happen?
- Add voice-over?
- Not clear who the man clicking yes/no is
- Improvements not clearly connected to risks

#### **P6**

- Add voice-over?
- Changing social values is not obvious from the video
- Make problems more clear
- Was wondering if these problems are connected?
- Make a clearer separation between problem and intervention – visually distinguishing them more
- Is contestation meant to be a one-time thing, because that is what the video now shows, or should it be a sustained thing?
- They expected people contesting with system directly, rather than humans operating the system

#### **P7**

- The AI side of value change wasn't brought up
- Not clear who the human in the loop was, thought it maybe was a citizen, and part of the contestation
- Appeals show a human and a human, no involvement of AI

- Appeals doesn't feel "system based" but more like the way it's currently done
- The adversarial dialogue shown does not involve the AI
- Show how the camera car sees the world. This might fix everything.
- Contestation was unclear. Who are the people involved?