

# SceneFinder - A Game that Elicits Tacit Knowledge from Crowds

How can we collect common-sense knowledge about scenes and rooms for machine learning models by using a collaborative, multiplayer game?

## 1. Background

- Machine learning models require large amounts of data
- Tacit knowledge: Intuitive, common-sense knowledge
- Explicit knowledge: Information available online/ databases
- Collecting tacit data is expensive and time consuming

## 2. Method

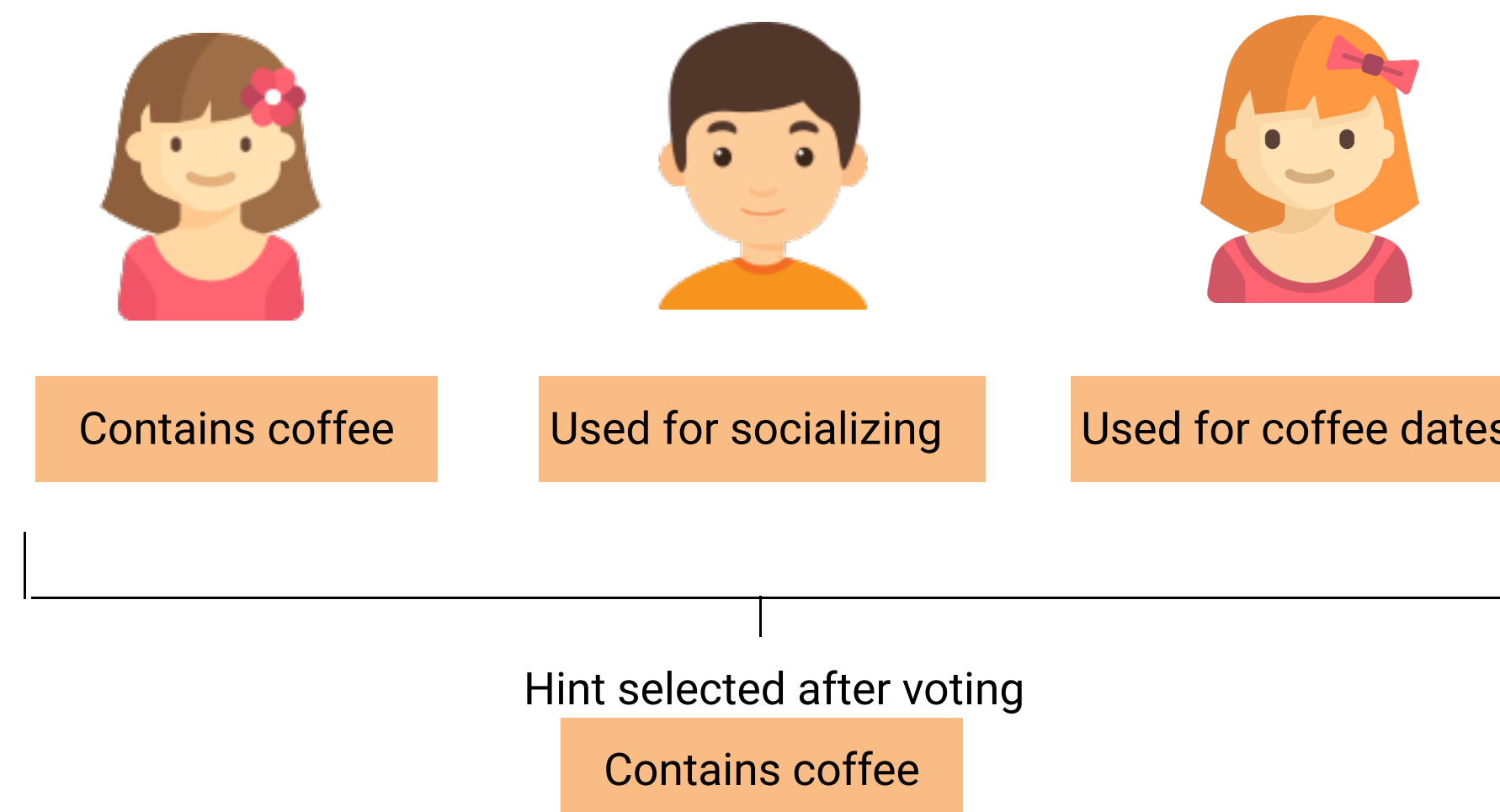
- Collect information about how tacit knowledge has been collected in the past
- Design a game capable of collecting diverse, reliable data
- Benchmark the quality of collected data to evaluate the performance of the game

## 3. Game Design

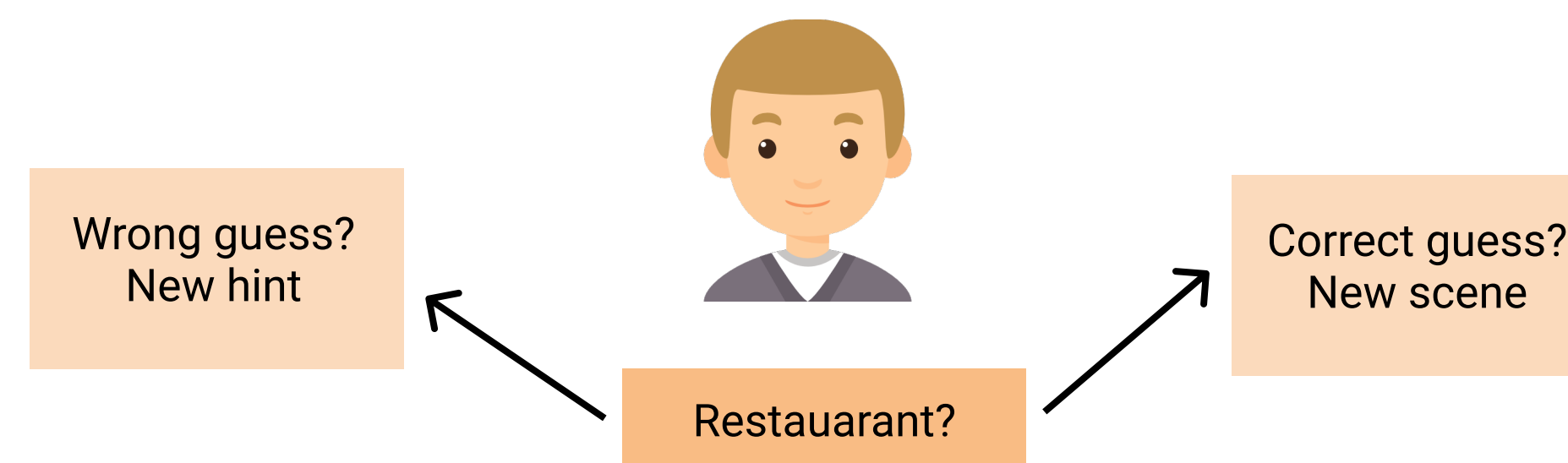
- Guessing game where 3 narrators work together to help a guesser guess the name of scene based by suggesting hints
- Each narrator suggests a hint
- The narrators vote for the hint they think is the most relevant which is sent to the guesser
- The guesser makes a guess based on this most relevant hint

## 4. Game flow example for 'cafe' scene

Narrators give hints and vote for most relevant one



Guesser uses this hint to try and guess the correct scene



Icon by Raj De on freeicons.io    Freepik from www.flaticon.com

## 5. Data Collection

- Each hint gets points based on the number of votes
- Aggregating votes over a set of games for the same scene results in a collection of ranked hints that can be used to classify the scene
- Use of premade hint templates helps us control the context of hints

## 6. Results

- 247 facts collected about 12 scenes in 3 games.
- 98% facts were true
- Relevance based ranking of hints obtained
- High game engagement

## 7. Conclusion

- Data collected was accurate and reliable
- Data lacked diversity, and game favored explicit knowledge over tacit, can be fixed by introducing taboo words in the future

Supervisors: Agathe Balayn, Ujwal Gadiraju, Jie Yang