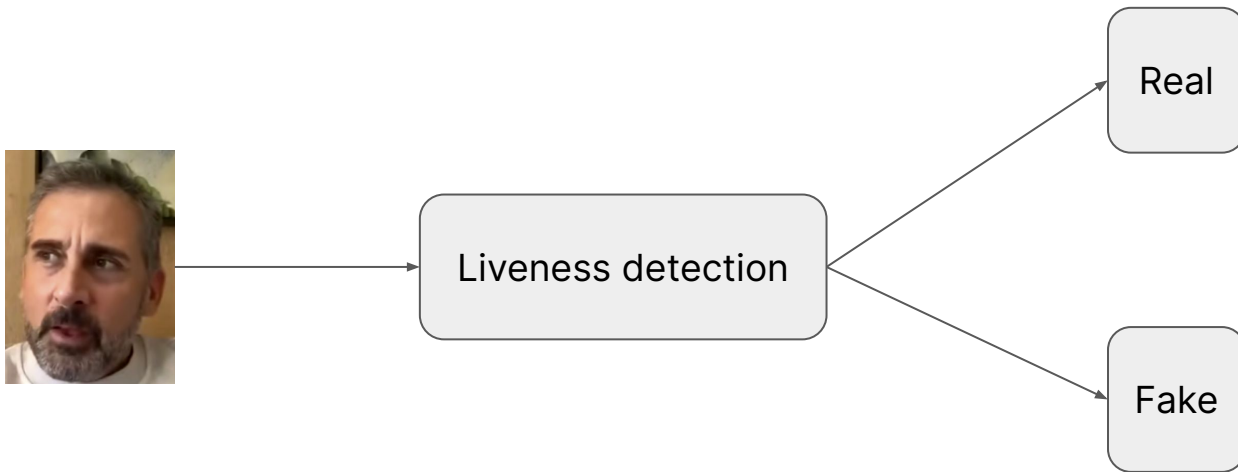


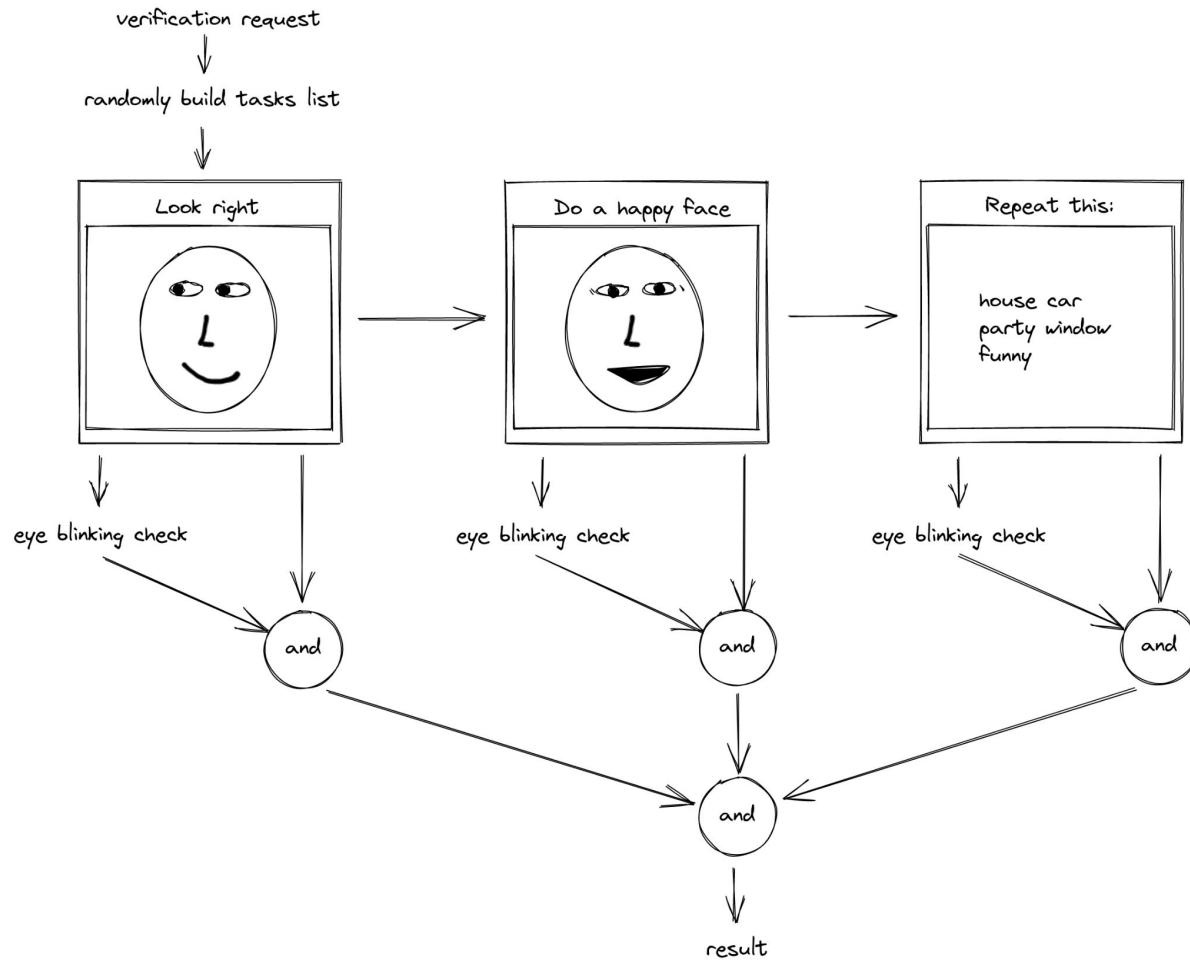
LittleAntispoof

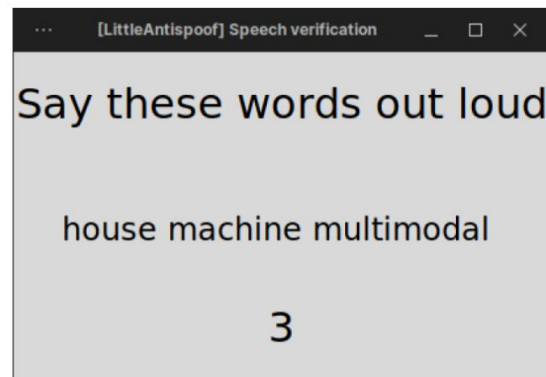
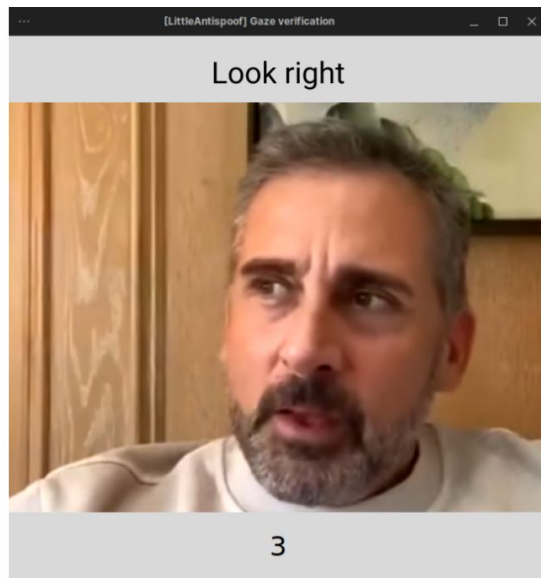
A multimodal face liveness detection system

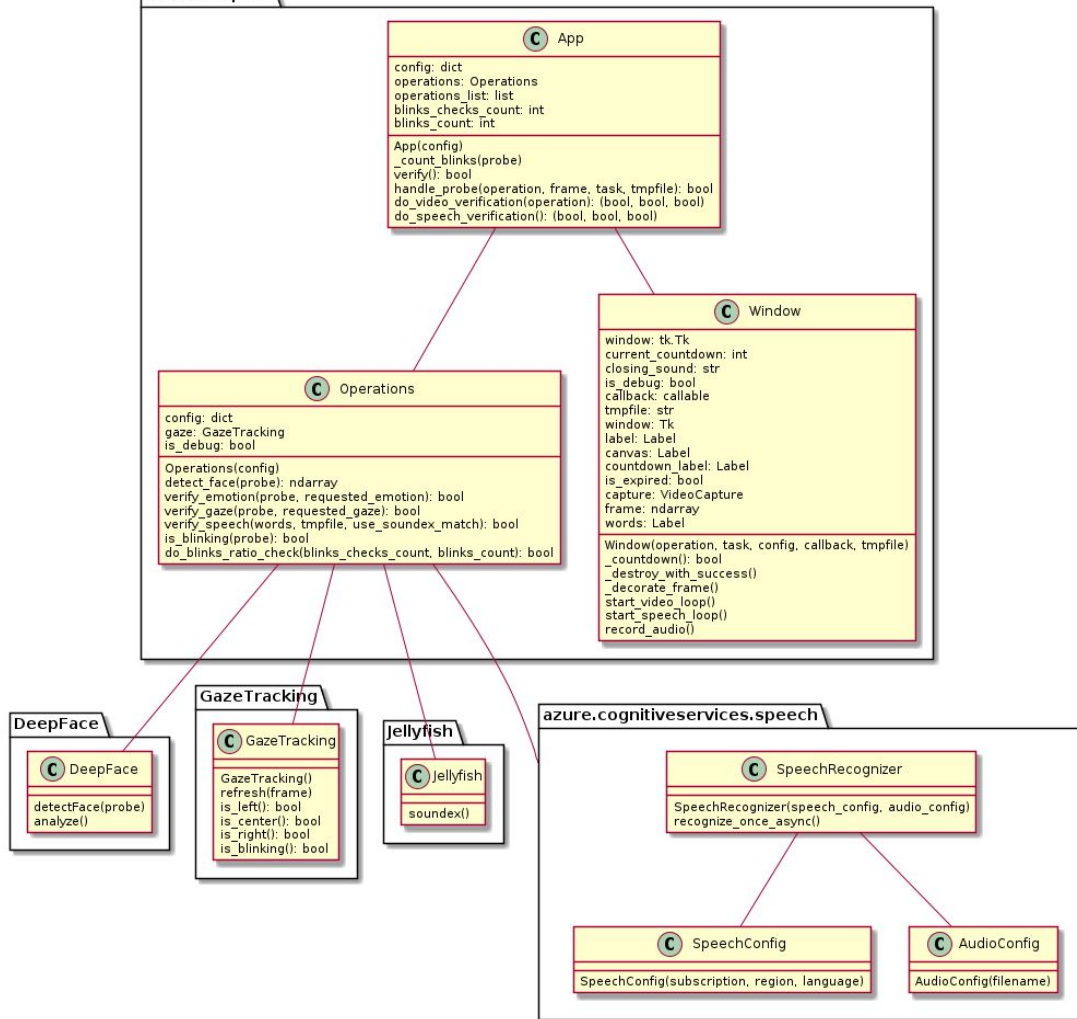
Davide Quaranta, Federico Cernera

What is liveness detection?









System structure

Additional utility libraries:

- OpenCV
- PIL
- playsound
- sounddevice
- python-dotenv

DeepFace: emotion recognition



DeepFace.analyze()

```
{
  'emotion': {
    'angry': 2.0189900733421686e-06,
    'disgust': 6.294248737258371e-14,
    'fear': 5.569503523939416e-08,
    'happy': 99.99113678879361,
    'sad': 8.27286306586789e-06,
    'surprise': 7.136712091459289e-07,
    'neutral': 0.008858164601887143
  },
  'dominant_emotion': 'happy',
  'region': {
    'x': 245,
    'y': 76,
    'w': 221,
    'h': 221
  }
}
```

GazeTracking: gaze direction detection



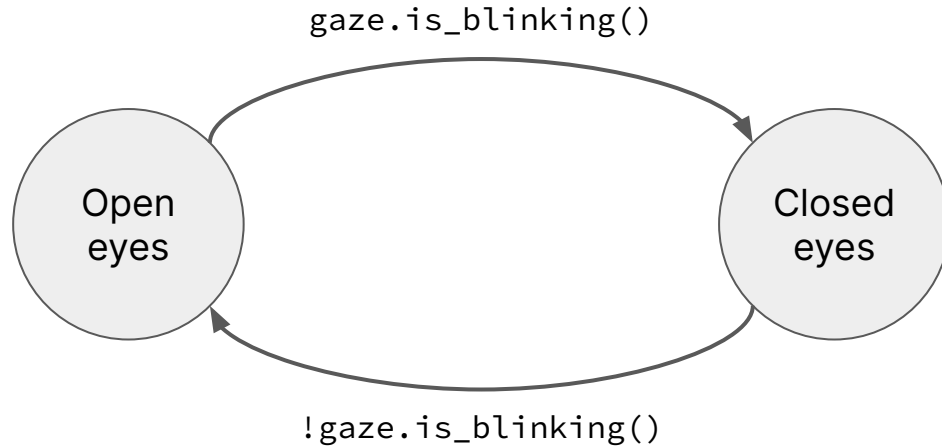
GazeTracking

Left

Center

Right

GazeTracking: blinking detection

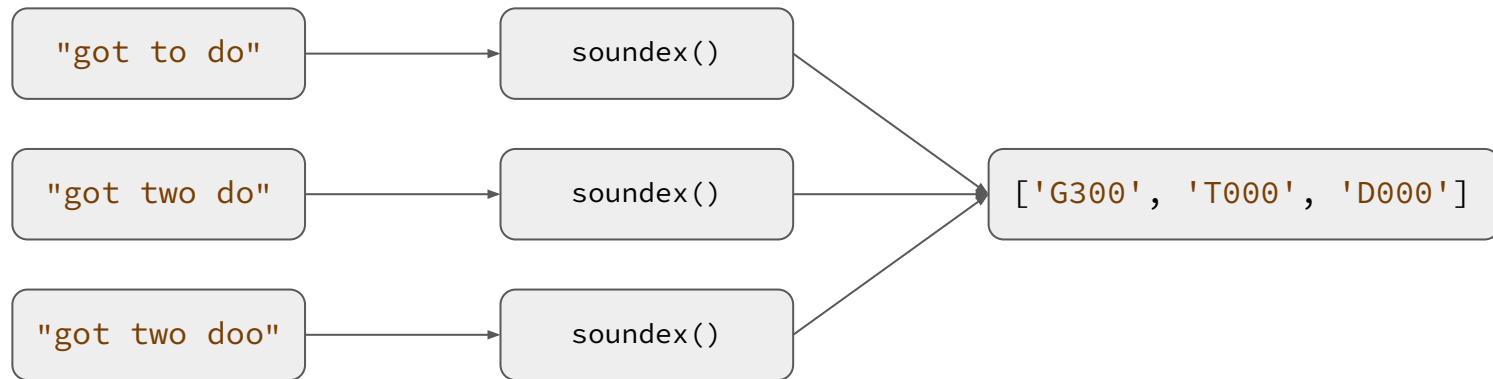


Azure STT: speech recognition



```
SpeechRecognitionResult(  
  result_id=baf1c09d86664b868e7e564eadebbfc3,  
  text="the user said this text here",  
  reason=ResultReason.RecognizedSpeech  
)
```

Jellyfish: phonetic matching

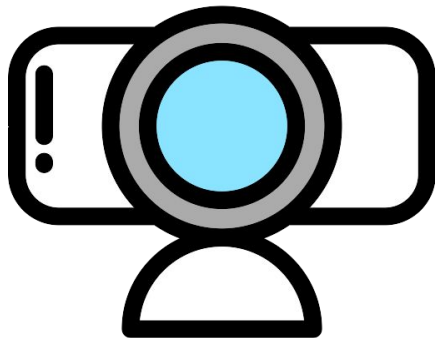


System design: configurability

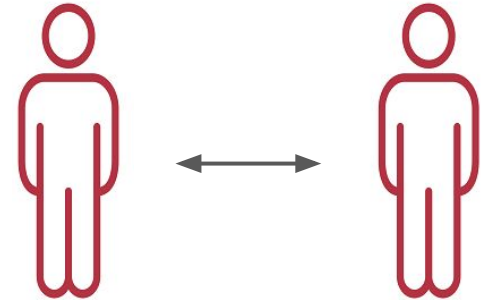


```
{
  "face_detector": "opencv",
  "window_duration_secs": 5,
  "blinking": {
    "min_threshold": 0.2,
    "max_threshold": 0.6
  },
  "speech": {
    "words_to_display": 3,
    "use_soundex_match": true,
    "record_duration_secs": 5
  },
  "sounds": {
    "start": "./sounds/start.mp3",
    "camera_shot": "./sounds/camera_shot.wav",
    "mic_stop": "./sounds/mic_stop.wav"
  },
  "debug": true
}
```

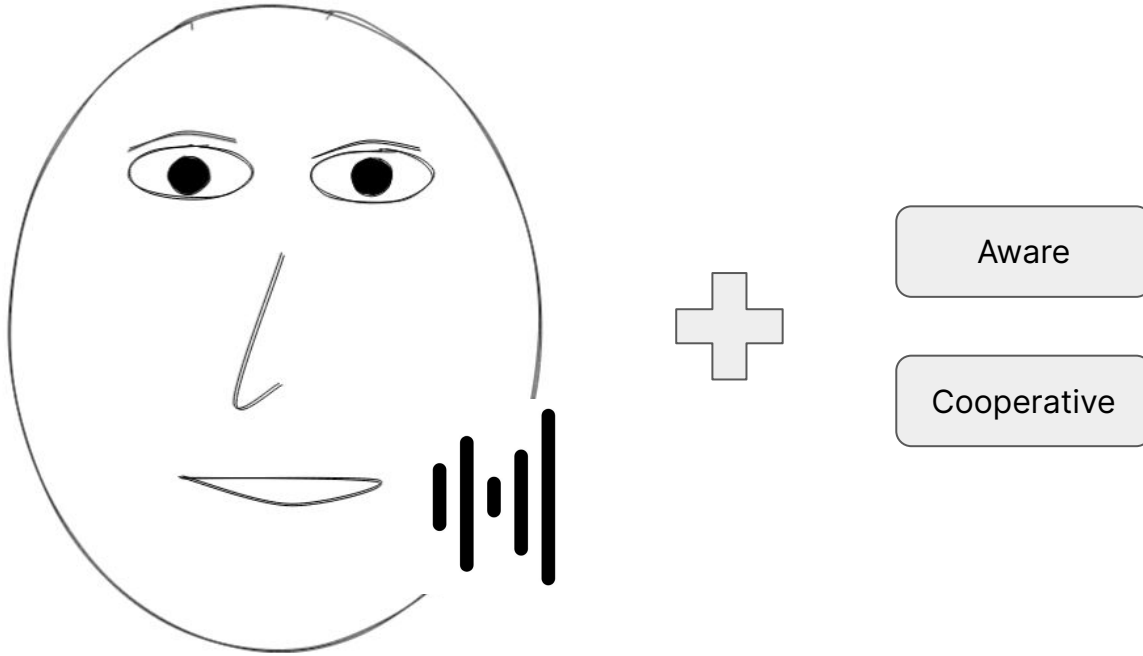
Context: computational environment



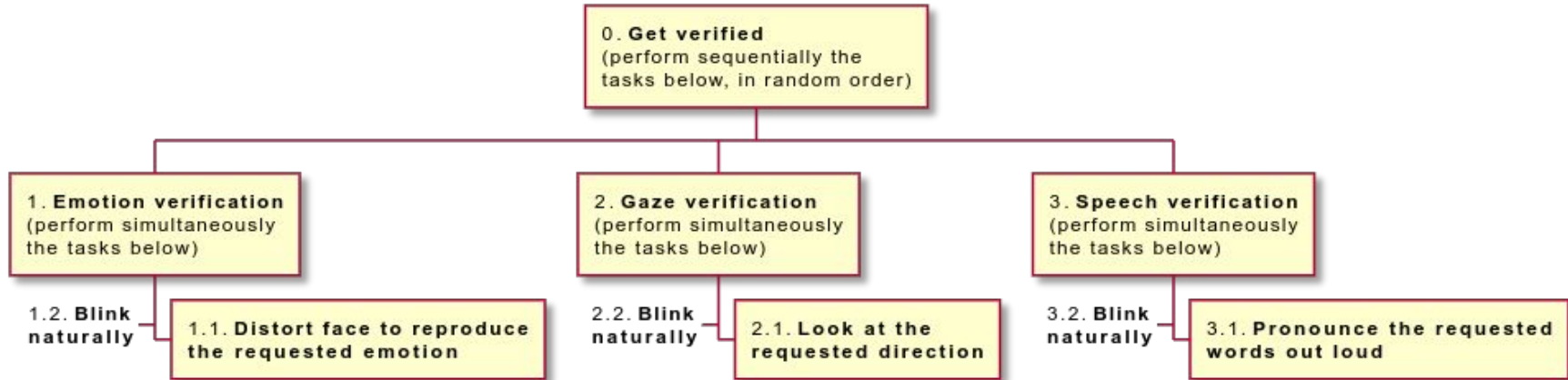
Context: physical environment



Context: user environment



Hierarchical Task Analysis

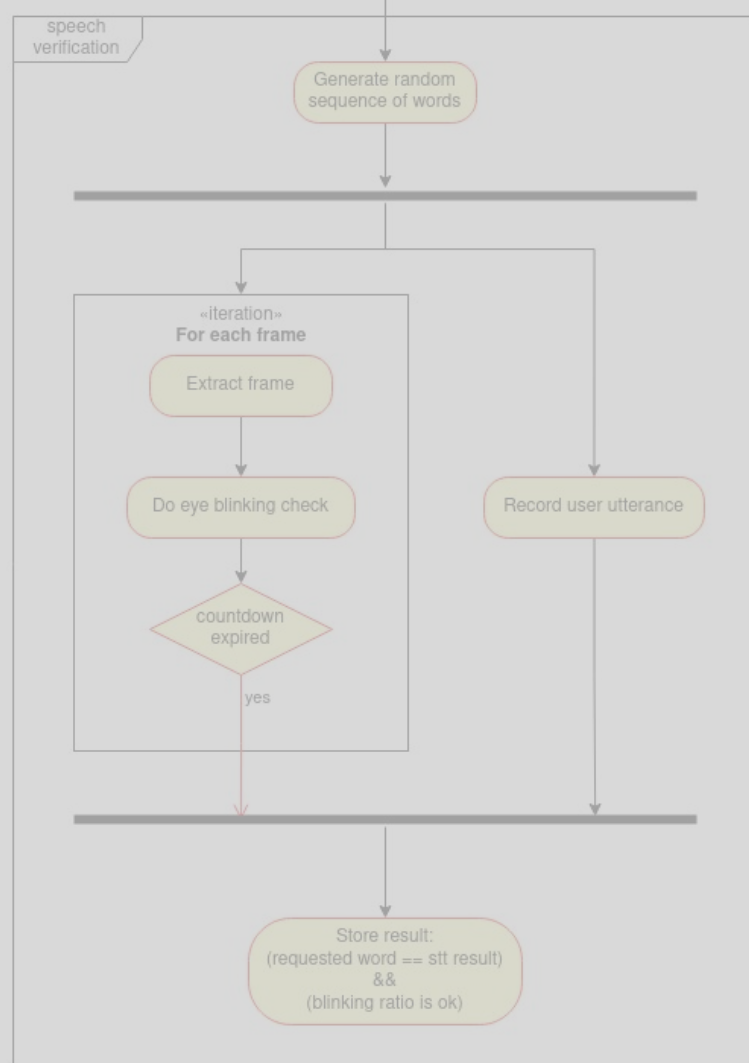
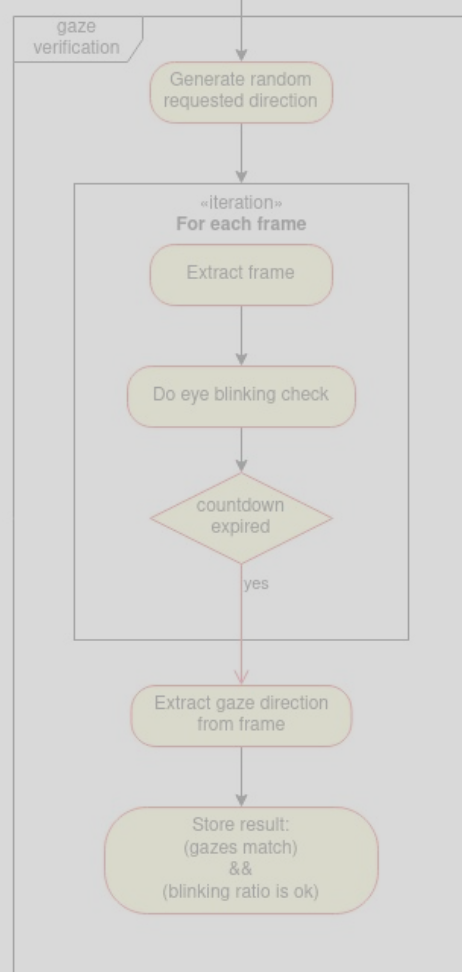
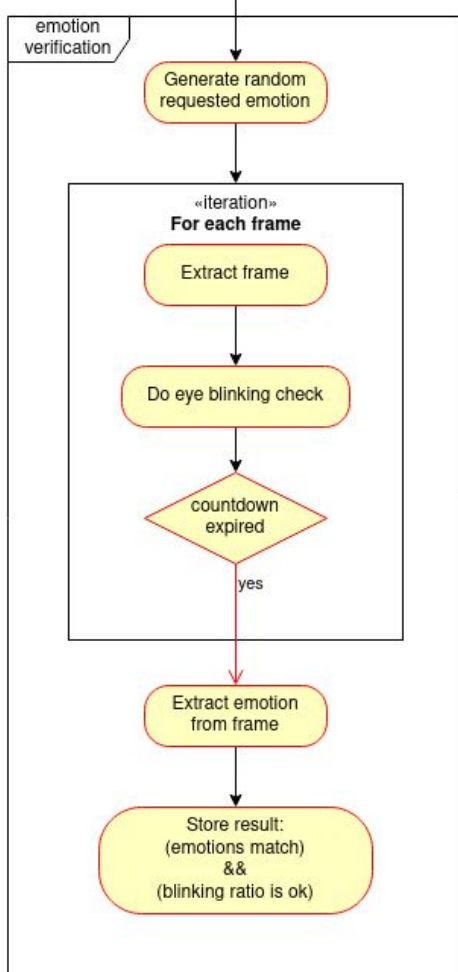


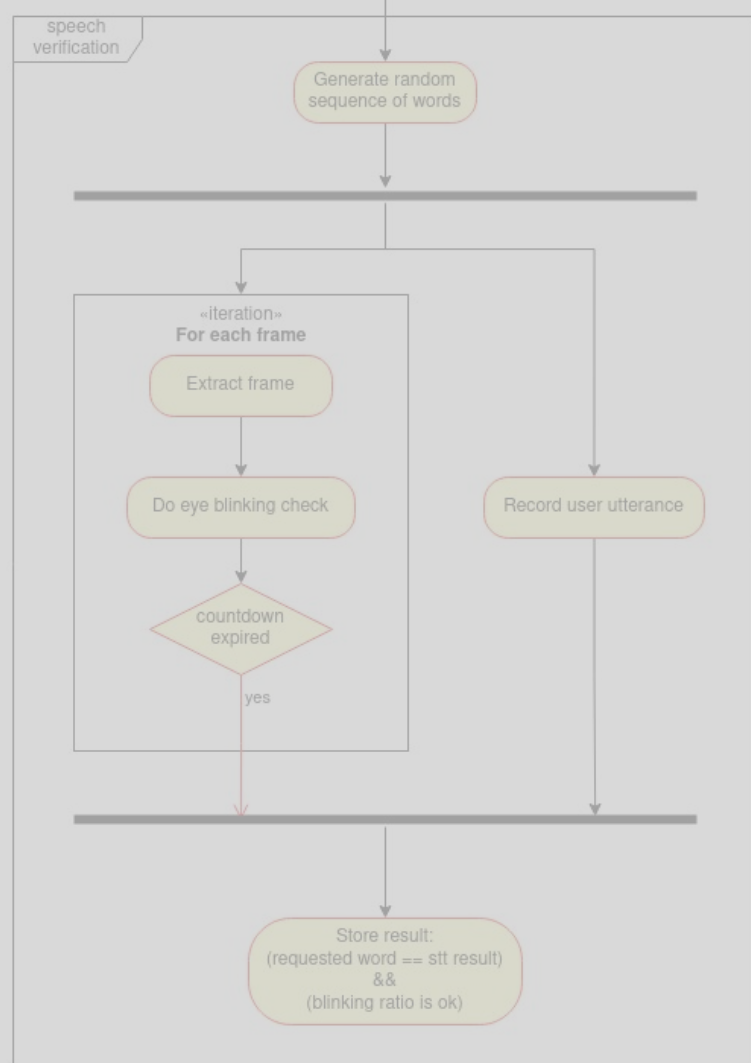
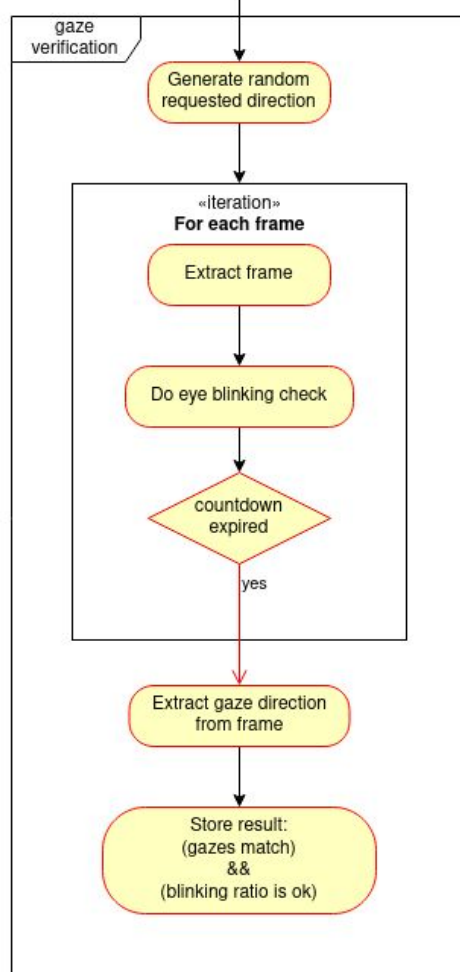
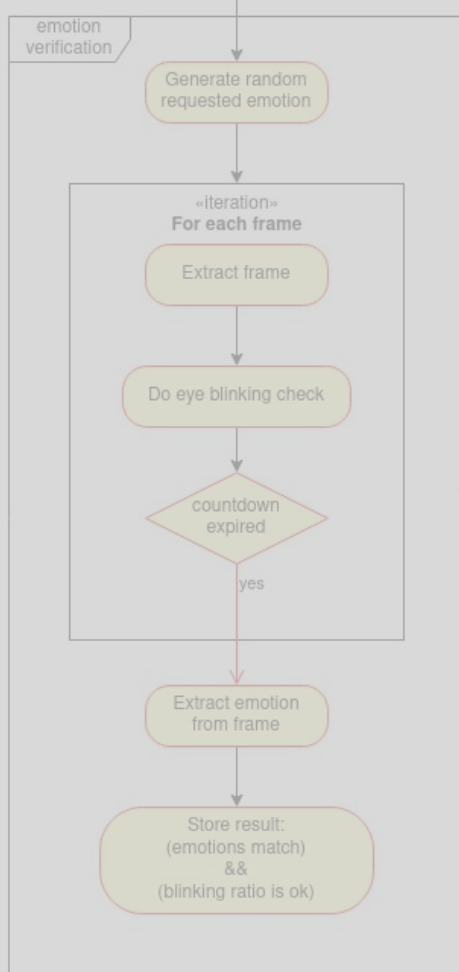
Multimodal fusion

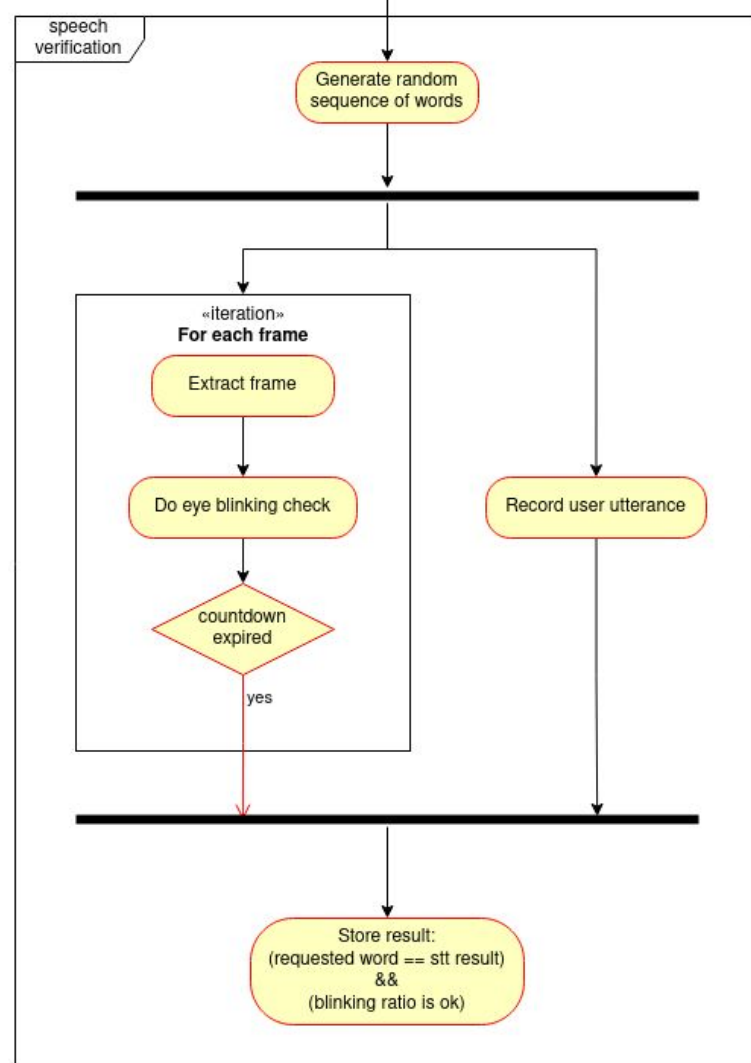
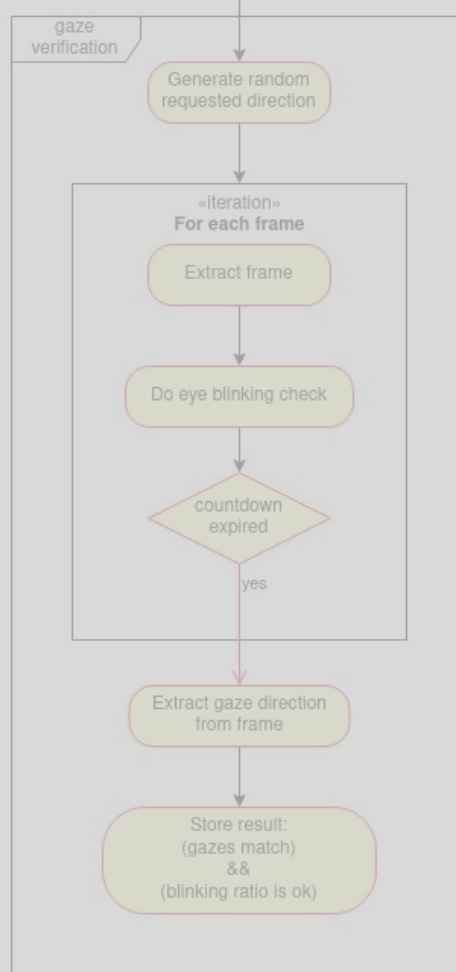
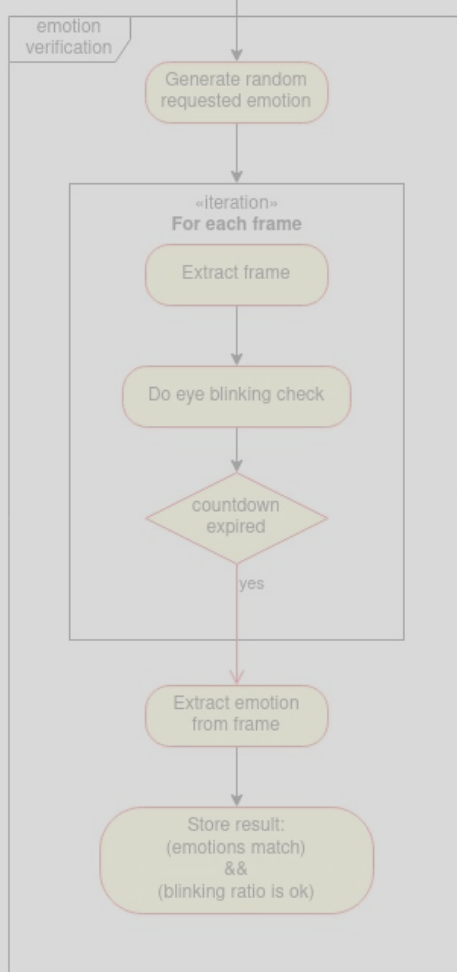
Information type	Modality
Emotion	Gestural
Gaze	Gestural
Speech	Speech
Blinking	Gestural (implicit)



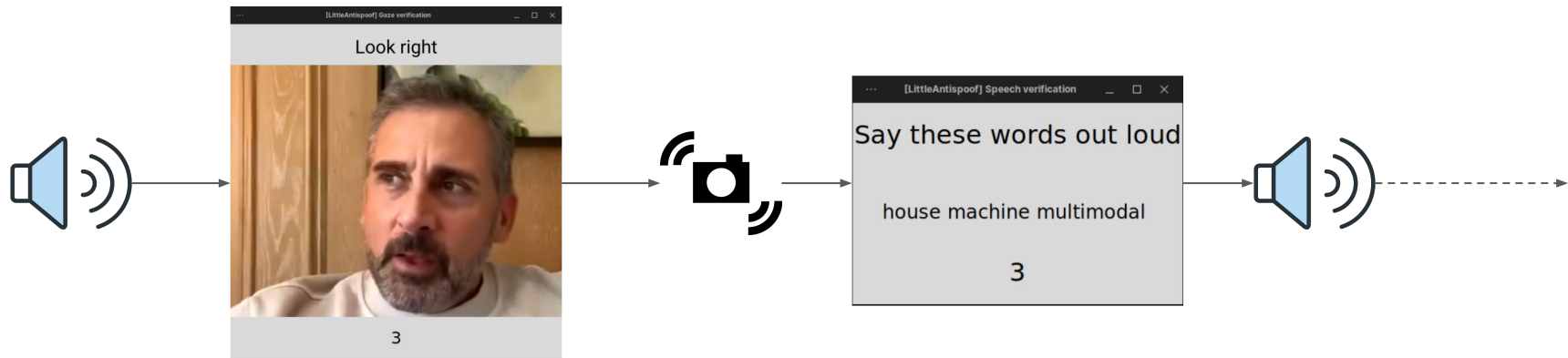
Decision-level fusion







Multimodal fission



Conclusion

- Face **liveness detection** module.
- Challenge-response style.
- **Emotion** recognition.
- **Gaze** detection.
- **Speech** recognition.
- **Phonetic matching**.
- **Blinking** check.

