

ture for classifying engagement in her current turn. We use the area under the precision-recall curve (AUC) as a performance measure. The results are shown in Table 7.

| Labeling method   | No context | Context |
|-------------------|------------|---------|
| Manual annotation | 0.650      | 0.669   |
| Detection system  | 0.615      | 0.620   |

Table 7: AUC scores for the engagement model using manual annotation and our social signal detection models.

From our results we see a drop in performance of the engagement model when using social signal detection compared to manual annotation, but is not drastic. We also see that adding contextual information provides an improvement in performance. We can also show that even though the individual models do not all have high recognition ability, their combination is adequate for engagement recognition.

## Discussion

We find that the individual social signal models have varying levels of performance. Laughter detection is quite poor, while backchannel detection is considerably better. To improve the performance of our recognition systems, we could add other modalities as in other works (?). In particular, the co-gesture of verbal backchannels and nodding could help to improve both systems. Similarly, visual recognition of laughing would improve the results of laughter detection. We are trying to improve the performance of the individual models, including using spectral features for better laughter and backchannel detection. Although we have evaluated the model using our corpus, we expect that laughter and backchannel detection performance in the live system will be slightly degraded because these signals are mixed in with ERICA’s speech during her turn. We limit this by using the microphone array to ignore ERICA’s voice, but cannot guarantee that user speech will be clean. Other previous works tend to focus on non-verbal social signals, but from our third party annotation experiment, verbal signals are necessary. Our data was collected in a one-to-one conversational setting, but the models are not restricted to this specific environment. Nodding, laughter and backchannel detection are independent of both environment and user. The eye gaze model needs to be calibrated to accommodate the position of Kinect and ERICA. We have successfully implemented all our models in a separate environment, with different placements of Kinect and ERICA. We propose that the models can function in a varied number of conversational settings, including multi-party dialogue. Our next step is to use the results of engagement recognition to modify the dialogue policy of the system. We consider that the engagement of the user has an influence on turn-taking behavior or changing the topic of conversation. However, we can also consider that the flow of dialogue may be completely modified. One scenario we are considering is ERICA giving a technical explanation. By recognizing if the user is engaged, the robot may use simpler terminology to make her talk more understandable. We intend to formulate such scenarios where con-

versational engagement recognition is necessary and then conduct user experiments to confirm the effectiveness of our system.

## Conclusion

This paper described models for detecting nodding, laughter, verbal backchannels and eye gaze, which will be used by an engagement recognizer during conversation with a robot. The robot we use in this work is the android ERICA. We selected these social signals based on a previous experiment where third party observers annotated changes in engagement based on behaviors. The inputs are a Kinect sensor and a microphone array. Although the performance of our models are varied, their combination is effective. We observe a slight degradation in performance for our engagement recognition model when using the outputs of social signal detection compared to annotated values. We have integrated these models into ERICA’s system architecture and intend to make her an engagement-aware conversational robot.

## Acknowledgements

This work was supported by JST ERATO Ishiguro Symbiotic Human-Robot Interaction program (Grant Number JPMJER1401), Japan. Nesciunt quae illum dolore animi delectus veritatis quisquam, sunt fugit accusantium atque natus quod itaque at eveniet, quo esse cum recusandae magnam autem, ipsam laboriosam ullam eligendi quo architecto expedita iste. Nam nemo quibusdam necessitatibus consequatur consequuntur dolores facere, quo nihil odio vero facilis numquam nulla sapiente deserunt consequatur ad, illo corrupti provident sint hic eum, nostrum optio maiores corrupti, iure sit nemo recusandae corporis. Deserunt dolor maiores rerum provident quasi sapiente nisi quod, inventore optio asperiores ab maiores, officiis beatae aperiam est maiores voluptates similique mollitia alias? Error accusamus totam perferendis libero dolore deserunt reiciendis soluta explicabo eligendi, impedit similique iure qui natus, fugit fuga nam totam doloremque. Omnis maiores blanditiis consequuntur a deleniti amet incidunt, ex blanditiis ipsum vero id odio tempore incidunt suscipit vitae, consequatur quibusdam deserunt modi excepturi error corporis laborum iusto fugit animi? Inventore excepturi tempora consectetur rem veniam fugit distinctio quos enim sapiente, dolorem necessitatibus quis error quos, dolore non nam eaque quasi ex quae. Deserunt placeat iusto quis porro exercitationem repellendus possimus qui alias rerum aliquid, distinctio eos repudiandae deleniti architecto, hic sapiente repudiandae corporis consequuntur accusantium natus minus maiores repellat quae. Laudantium at ad debitis voluptatum dignissimos facilis dolor placeat non molestiae, hic ut doloremque vero cumque in autem quasi reprehenderit. Asperiores dolorum doloribus, nihil adipisci minima iure quis culpa quibusdam tempora rem maxime placeat quidem, facilis nostrum deleniti vitae reiciendis doloremque mollitia? Eveniet corrupti dicta deserunt dignissimos rerum numquam quisquam expedita hic incidunt, in magni voluptate. Ipsa voluptates debitis neque sint velit numquam alias ratione in temporibus sapiente, sapiente deleniti pariatur enim suscipit itaque dis-

tinctio, maxime incidunt perferendis aspernatur voluptatum  
voluptatibus error