CATS2021: International Workshop on Corpora And Tools for Social skills annotation

BEATRICE BIANCARDI, LTCI, Télécom Paris, IP Paris, France
ELEONORA CECCALDI, DIBRIS, University of Genoa, Italy
CHLOÉ CLAVEL, LTCI, Télécom Paris, IP Paris, France
MATHIEU CHOLLET, LS2N, IMT Atlantique, France
TANVI DINKAR, LTCI, Télécom Paris, IP Paris, France

This Workshop aims at stimulating multi-disciplinary discussions about the challenges related to corpus creation and annotation for social skills behavior analysis. Contributions from computational, psychological and psychometrics perspectives, as well as applications including platforms to share corpora and annotations, are welcomed. The main challenges related to corpus creation include the choice of the best setup and sensors, finding a trade-off between eliciting natural interactions, limiting invasiveness and collecting precise information. The second issue in this context regards the process of annotation. The choice of the type of annotators (experts vs. nonexperts), the type of annotations (automatic vs. manual, continue vs. discrete), the temporal segmentation (windowed vs. holistic) is crucial for a correct measure of the phenomenon of interest and getting significant results. The topics of CATS2021 will have a strong impact on researchers and stakeholders across different disciplines, such as Computer Science, Social Signal Processing, Psychology, Statistics. Leveraging the opportunities offered by such a multidisciplinary environment, the participants could enrich their perspective, strengthen their practices and methodologies and draw together a research roadmap tackling the discussed challenges, which might be taken up in future collaborations.

CCS Concepts: • **Applied computing** → *Psychology*; • **General and reference** → *Empirical studies*; *Metrics*; *Experimentation*; • **Computing methodologies** → *Artificial intelligence.*

Additional Key Words and Phrases: Multi-modal Social Behavior Datasets; Multi-modal Behavior Analysis; Social Behavior Annotation; Social Signal Processing; Social Skills

ACM Reference Format:
Beatrice Biancardi, Eleonora Ceccaldi, Chloé Clavel, Mathieu Chollet, and Tanvi Dinkar. 2021. CATS2021: International Workshop on Corpora And Tools for Social skills annotation. In *Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21), October 18–22, 2021, Montréal, QC, Canada.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3462244.3480977

1 WORKSHOP MOTIVATION, EXPECTED OUTCOMES AND IMPACT

1.1 Motivation

Social skills represent a fundamental resource in any professional and personal situation for conducting smooth interactions [1]. Methods in Artificial Intelligence have become increasingly popular in order to automatically assess social skills – by analyzing multi-modal behavior in several contexts such as public speaking (e.g., [2, 3]), job interviews (e.g., [4, 5]), group interactions (e.g., [6, 7], in populations with dysfunctions [8], and involving both human-human and human-machine interactions [9]. These methods could facilitate planning interventions targeted at improving these

competencies, for example by giving appropriate feedback and personalized training [10]. The first step to investigate multi-modal cues of social skills, no matter the context, often consists in the analysis of corpora, by automatically extracting behavioral features and manually annotating more subjective and psychological constructs. Finding an existing corpus that could be exploited for one's research interest could be a difficult step, either because of difficulties in obtaining access, or because they are not adapted for different research goals. Often, researchers prefer to create their own corpus. This results in a large amount of existing corpora that are often not accessible to other researchers and are not fully exploited. The main challenges related to corpus creation include the choice of the best setup and sensors, finding a trade-off between eliciting natural interactions, limiting invasiveness and collecting precise information. The second issue in this context regards the process of annotation. The choice of the type of annotators (experts vs. non-experts), the type of annotations (automatic vs. manual, continue vs. discrete), the temporal segmentation (windowed vs. holistic) is crucial for a correct measure of the phenomenon of interest and getting significant results.

This Workshop aims at stimulating multi-disciplinary discussions about these challenges and sharing the best practices for analyzing social skills behavior. Contributions from computer science, psychological and psychometrics perspectives, as well as applications including platforms to share corpora and annotations, are welcomed.

1.2 Expected Outcomes

The expected outcome of the Workshop is to elaborate a state of the art of the most current issues that researchers face when conducting multi-modal social behavior analysis. Moreover, the Workshop is expected to foster new collaborations towards the creation of standards and practices that could facilitate the exploitation of existing resources around the topic.

1.3 Impact

The topics of the Workshop may have a strong impact on researchers and stakeholders across different disciplines, such as Computer Science, Social Signal Processing, Psychology, Statistics. Leveraging the opportunities offered by such a multidisciplinary environment, the participants could enrich their perspective, strengthen their practices and methodologies and draw together a research roadmap tackling the discussed challenges, which might be taken up in future collaborations.

2 WORKSHOP ORGANIZERS

Beatrice Biancardi (beatrice.biancardi@telecom-paris.fr) is a postdoctoral researcher at LTCI, Télécom Paris, Insitut Polytechnique de Paris, France. She obtained her PhD at Sorbonne University, under the supervision of Catherine Pelachaud. She currently works in the context of the AI4SoftSkills project, funded by the TSN and MINES Carnot institutes. Her research interests include social behavior modeling by applying social signal processing to human-human and human-agent interaction.

Eleonora Ceccaldi (eleonora.ceccaldi@edu.unige.it) is a professional psychologist and a PhD student in Computer Science at the University of Genoa (Italy). She works on event segmentation, event structure perception and food-related social interaction. She co-organized the first International Workshop on Multi-Scale Movement Technologies at the ICMI2020 conference. She co-authored several research papers. She is a Scientific Explainer for the Genoa Science Festival.

Chloé Clavel (chloe.clavel@telecom-paris.fr) is Professor in Affective Computing at Télécom-Paris, Institut Polytechnique de Paris, where she coordinates the Social Computing topic. Her research work deals with interactions between humans and virtual agents, from user's socio-emotional behavior analysis to socio-affective interaction strategies with a focus on speech and language processing.

Mathieu Chollet (mathieu.chollet@imt-atlantique.fr) is an Assistant Professor at IMT Atlantique and a member of the LS2N laboratory. His research deals with human communicative behavior modeling and virtual social interactions for social skills training.

Tanvi Dinkar (tanvi.dinkar@telecom-paris.fr) is a PhD student at Télécom Paris, Insitut polytechnique de Paris, under the supervision of Prof. Chloé Clavel. She works on representations of spontaneous speech phenomena (such as disfluencies) to reflect metacognitive states. Her research interests include spoken language understanding, psycholinguistics, and communicative strategies.

3 WORKSHOP FORMAT

The target duration is a full day. The Workshop will host both invited presentations and contributions upon submission in response to a call for papers. People are welcomed to attend the Workshop even without submitting any contribution. Contributions will include both short and long papers, following ICMI format. We received 10 contributions coming from different research areas.

The Workshop is divided into 2 sessions. The first session is about the challenges relative to corpus creation. The second session addresses challenges related to the annotation process. A round table, animated by the Workshop organizers, is scheduled at the end of the day, with the goal to summarize the discussions during the Workshop, comment on them, and help identify future directions.

3.1 Keynote Speakers

- Daniel Gatica-Perez, Professor, Idiap, EPFL, Switzerland
- Laura Cabrera-Quiros, Assistant Professor, Costa Rican Institute of Technology
- Tobias Baur, Research Associate, Chair for Human-Centered Artificial Intelligence, Augsburg University

3.2 Areas of Interest

Topics centered around the challenges related to social skills annotation of multi-modal behavior analysis, in the context of human-human, human-agent or human-robot interaction, including but not limited to:

- New Multi-modal corpora for studying social skills
- Review of existing work about social skills analysis
- Novel techniques to extract and annotate verbal and non-verbal behavior
- Annotation of subjective constructs related to social skills
- Annotation Tools
- Integration of existing corpora with annotations for social skills analysis
- Annotation schemes
- Data transformation and manipulation

- Data representation
- Platforms for storing and sharing corpora

3.3 Scientific Committee

A panel of experts from different disciplines agreed to participate in the Scientific Committee and to contribute to the peer-review process. The peer-review is double-blind and at least two reviewers are assigned to each paper.

We thank the Scientific Committee members:

Lucien Maman, LTCI, Télécom Paris, France
Mireille Fares, CNRS-ISIR, Sorbonne University, France
Brian Ravenet, LISN-CNRS, Paris-Saclay University, France
Giovanna Varni, LTCI, Télécom Paris, France
Catherine Pelachaud, CNRS-ISIR, Sorbonne University, France
Magalie Ochs, CNRS-LIS, Aix Marseille University, France
Fabien Boucaud, Heudiasyc, UTC of Compiègne, France
Dinesh Babu Jayagopi, IIIT Bangalore, India
Alexander Heimerl, Augsburg University, Germany
Mohammad Rafayet Ali, Department of Computer Science, University of Rochester, United States
Jean-Claude Martin, LISN-CNRS, Paris-Saclay University, France
Nicu Sebe, DISI, University of Trento, Italy
Hayley Hung, Perceptive Computing Lab, TU Delft, Netherlands
Cigdem Beyan, Italian Institute of Technology, Italy
Hiroki Tanaka, Nara University, Japan
Radoslaw Niewiadomski, University of Trento, Italy
Bernd Dudzik, Perceptive Computing Lab, TU Delft, Netherlands
Chirag Raman, Perceptive Computing Lab, TU Delft, Netherlands

REFERENCES

[1] G. Sharma and P. Sharma, "Importance of soft skills development in 21st century curriculum.," *International Journal of Education & Allied Sciences*, vol. 2, no. 2, 2010.
[2] L. Chen, G. Feng, J. Joe, C. W. Leong, C. Kitchen, and C. M. Lee, "Towards automated assessment of public speaking skills using multimodal cues," in *Proceedings of the 16th International Conference on Multimodal Interaction*, pp. 200–203, 2014.
[3] M. Chollet and S. Scherer, "Assessing public speaking ability from thin slices of behavior," in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 310–316, IEEE, 2017.
[4] L. S. Nguyen, D. Frauendorfer, M. S. Mast, and D. Gatica-Perez, "Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior," *IEEE transactions on multimedia*, vol. 16, no. 4, pp. 1018–1031, 2014.
[5] L. Hemamou, G. Felhi, V. Vandenbussche, J.-C. Martin, and C. Clavel, "Hirenet: A hierarchical attention model for the automatic analysis of asynchronous video job interviews," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 573–581, 2019.
[6] D. Sanchez-Cortes, O. Aran, D. B. Jayagopi, M. S. Mast, and D. Gatica-Perez, "Emergent leaders through looking and speaking: from audio-visual data to multimodal recognition," *Journal on Multimodal User Interfaces*, vol. 7, no. 1, pp. 39–53, 2013.
[7] B. Biancardi, L. Maisonnave-Couterou, P. Renault, B. Ravenet, M. Mancini, and G. Varni, "The wonowa dataset: Investigating the transactive memory system in small group interactions," in *Proceedings of the 2020 International Conference on Multimodal Interaction*, pp. 528–537, 2020.
[8] C. Bloch, K. Vogeley, A. L. Georgescu, and C. M. Falter-Wagner, "Intrapersonal synchrony as constituent of interpersonal synchrony and its relevance for autism spectrum disorder," *Frontiers in Robotics and AI*, vol. 6, p. 73, 2019.

[9] S. M. Anzalone, J. Xavier, S. Boucenna, L. Billeci, A. Narzisi, F. Muratori, D. Cohen, and M. Chetouani, "Quantifying patterns of joint attention during human-robot interactions: An application for autism spectrum disorder assessment," *Pattern Recognition Letters*, vol. 118, pp. 42–50, 2019.

[10] L. Batrinca, G. Stratou, A. Shapiro, L.-P. Morency, and S. Scherer, "Cicero-towards a multimodal virtual audience platform for public speaking training," in *International workshop on intelligent virtual agents*, pp. 116–128, Springer, 2013.