

Ethics for Crowdsourced Corpus Collection, Data Annotation and its Application in the Web-based Game iHEARu-PLAY

Simone Hantke^{1,2}, Anton Batliner^{1,3}, and Björn Schuller^{1,4}

¹ Chair of Complex & Intelligent Systems, University of Passau, Germany

² Machine Intelligence & Signal Processing Group, Technische Universität München, Germany

³ Pattern Recognition Lab, FAU Erlangen-Nürnberg, Germany

⁴ Department of Computing, Imperial College London, UK

Email: simone.hantke@tum.de

Abstract

We address ethical considerations concerning iHEARu-PLAY, a web-based, crowdsourced, multiplayer game for large-scale, real-life corpus collection and multi-label, holistic data annotation for advanced paralinguistic tasks. While playing the game, users are recorded or perform labelling tasks, compete with other players, and are rewarded with scores and different prizes. Players will have fun playing the game and at the same time support science. With this modular, cross-platform crowdsourcing game, different ethical and privacy issues arise. A closer look is taken on ethics in recording of private content, data collection, data annotation, and storage, as well as sharing the data within iHEARu-PLAY. Further, we address the interplay of science and society in ethics and relate this with our application iHEARu-PLAY.

Keywords: Crowdsourcing, Corpus Collection, Data Annotation, Ethics, iHEARu-PLAY

1. Introduction

Crowdsourcing – the process of distributing tasks to an open, unspecified group of people via the internet – is an arising collaborative approach in the area of speech and language processing; it can be harnessed for many different types of applications and offers instantaneously access to populations with specific knowledge and skills, everywhere on the globe, and for any spoken language. For annotating speech data, many projects employed crowdsourcing to save costs compared to expert human annotation in labs (Burkhardt et al., 2010; Zhai et al., 2013). Most crowdsourcing services rely on so called ‘click-workers’, which are being paid a rather low compensation for their work. Their jobs are often not very appealing and thus, their intrinsic motivation will be low. As is the case for most newly developed techniques, crowdsourcing also raises both hopes and doubts, certainties and also many questions. Eskenazi et al. (2013) give a general analysis of crowdsourcing for speech processing.

When dealing with crowdsourcing, many economic and ethical problems arise, which are related to the type of crowdsourcing services, the task to be addressed, the country where the click-workers perform the tasks, and the pertinent labour laws. Ethics is often equated with decisions of high moral magnitude and associated with weighty concepts of right and wrong. Although the relevance of ethics to daily experience is not always easy to assess, Seedhouse (1998) proposes a definition, highlighting this daily relevance by referring to ethics as a process of deliberation about “how best to conduct one’s life in the presence of other lives” – in fact, a sort of reformulation of Kant’s categorical imperative: “Act only according to that maxim whereby you can, at the same time, will that it should become a universal law.”. In fact, this simple statement refers to a very complex multivariate problem, and solutions should be found to deal with crowdsourcing services

in a more efficient and ethical way.

This paper presents an alternative way to collect annotated or recorded data by crowdsourcing non-professional subjects with the fun gaming platform iHEARu-PLAY (Hantke et al., 2015). iHEARu-PLAY motivates people by giving them a playful environment, where they can have fun and at the same time, voluntarily help scientific research projects by annotating data or recording prompted tasks. Instead of offering a financial incentive, people are primarily motivated to participate due to the joyful experiences of a game. Usually, the motivation for an individual to voluntarily contribute to a crowdsourcing project ranges from altruism, over ego, to a shared sense of purpose; yet, the pursuit of fun and enjoyment through games is also seen as an emerging trend (Good and Su, 2011). On top of the intrinsic motivation of playing a game and helping science, various prizes and awards can also be given, e. g., for the best scores, the most frequent users, and/or for randomly selected winners.

The present topic belongs to the broader field of Computational Paralinguistics (Schuller and Batliner, 2014; Batliner and Schuller, 2014). It seems that studies addressing ethics in connection with crowdsourcing so far dealt with paid work and not with the kind of voluntary work we are aiming at, cf. Silberman et al. (2010) and Schmidt (2013); Adda and Mariani (2013) address economic, legal and ethical aspects with crowdsourcing for speech, Adda et al. (2014) crowdsourcing in the context of big data.

Section 2 describes iHEARu-PLAY’s concept and main idea. The ethical and personal issues are addressed in Section 3, including the types of data collected and how it will be used. In Section 4, we deal with general considerations, before summarising this paper in Section 5.

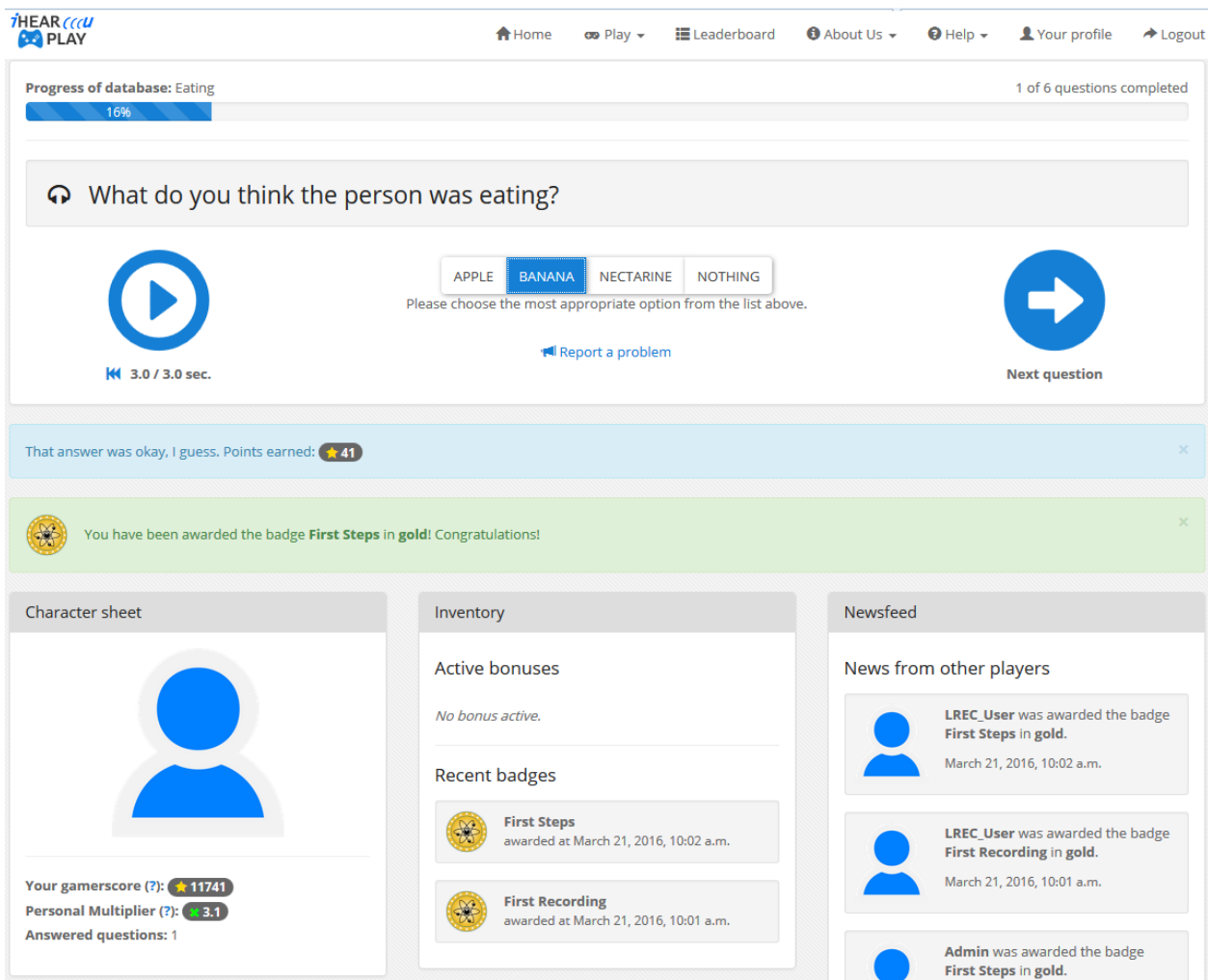


Figure 1: Exemplary screenshot: iHEARu-PLAY’s web-interface for labelling tasks, showing the progress for the database, list of answers, the feedback message after submitting the answer, including the earned points, followed by the next audio file and question.

2. Our Work – iHEARu-PLAY

iHEARu-PLAY¹ is a web-based multiplayer game for crowdsourced database collection and data labelling (Hantke et al., 2015); its primary purpose is holistic, multi-label annotation of multi-modal affective speech databases usually containing also or only speech. Existing speech and video databases, and also image databases can be added, and labelling tasks can be defined via a web-interface. Further, new speech data can be collected by players performing prompted recording tasks in the wild. Players perform these labellings or prompted recording tasks and are rewarded with prizes and scores based on the ‘correctness’ of their annotations, e. g., the agreement with a pre-defined gold standard (an already existing annotation from a former lab annotation task) or the agreement with the (majority vote of the) other players.

When a new user visits iHEARu-PLAY for the first time, (s)he will be able to access a demonstration that explains the idea behind iHEARu-PLAY and teaches the user how to interact with the system. After signing up at iHEARu-PLAY, the player can choose a database for the annotation

or recording task. Having picked an annotation database, the user will be presented with a random audio file and a question from that database. After playback of half of the audio file, a list of answers fades in from which the user can select one (or sometimes multiple). Having selected his or her answer, a submit button will fade in, which, upon execution, immediately presents a feedback message (based on the player’s performance), including the earned points as well as the next audio file and question. Then, the whole process starts over again. If the user earned a badge, it will be displayed in the same area as the feedback message and – if the user allowed to share his or her activity on the platform – on the activity ticker, thus visible to all other players. Figure 1 shows the web-interface for such a labelling task. The web-interface for the recording task is build up similarly as the labelling web-interface and just differs from it in the small substituted part shown in Figure 2. Users can start the recording by clicking on the microphone, read the above shown sentence out loud, and get presented with a live spectrogram visualising their recorded speech.

After the user stopped the recording, (s)he can listen to the recorded prompt, if wished record the same sentence again and finally upload the recorded prompt to iHEARu-PLAY.

¹<https://ihearuplay.fim.uni-passau.de/>

🎤 The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak.

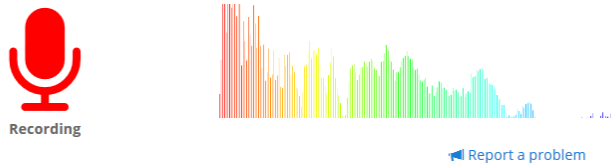


Figure 2: Exemplary screenshot: iHEARu-PLAY’s web-interface for recording tasks, showing the microphone to start and stop the recording and the spectrogram to visualise the recorded speech of the user.

🎤 The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak.

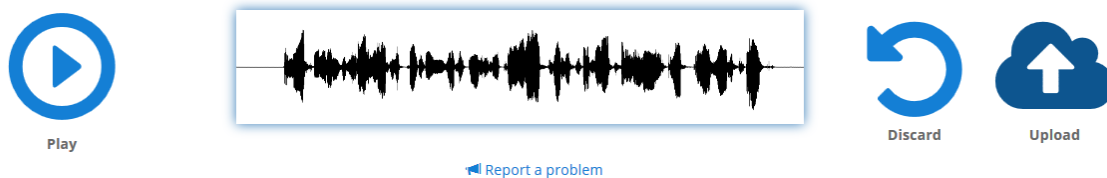


Figure 3: Exemplary screenshot: iHEARu-PLAY’s web-interface for recording tasks after having recorded a sentence, showing the possibility to listen to the recorded audio file again, discard it or upload it to iHEARu-PLAY, followed by the next sentence.

Figure 3 shows the recording part of the web-interface. iHEARu-PLAY is implemented with the open source high-level PythonWeb framework Django (Foundation, 2014 Version 17) and can be installed on Unix and Windows platforms. Its modular architecture allows for easy integration of custom extensions: New gaming components can be added as plugins in order to support new databases and modalities. The game will be available to the research community as a ready-to-use web-service. Researchers can add their own databases, optionally post rewards, and receive annotation results in the end. General users can register to play the game, record new data or annotate already existing data, gain points for the tasks performed, compete with other players on the leaderboard, have fun, and at the same time support science (Hantke et al., 2015).

3. Privacy and Ethical Issues

An absolutely fundamental step is to determine the ‘status’ of users in a participant agreement. Many crowdsourcing platforms include in their terms of use a statement that defines the workers as ‘independent contractors’. For instance, the Amazon Mechanical Turk Participation Agreement contains the statement: “As a Provider, you are performing Services for a Requester in your personal capacity as an independent contractor and not as an employee of the Requester.” (Turk, 2016). In contrast to the typical crowdsourcing platforms, iHEARu-PLAY does not define the users as independent contractors but as volunteers. There is no monetary compensation for the tasks per-

formed; iHEARu-PLAY is free of charge and only asks for voluntary participation.

For research scenarios, where data are collected, participating volunteers have to give informed consent. The iHEARu-PLAY informed consent form is included as an appendix to its platform. To ensure data anonymity and security, iHEARu-PLAY gives different restrictions and prohibitions within this form for users, e. g., an age restriction, the prohibition to give away personal information of themselves or other persons within the recordings or anywhere else on the platform, or generate unethical or inappropriate data (e. g., issues related to sexual or propaganda content). Within iHEARu-PLAY, there is a possibility for users to report other users if they generate or publish data against the privacy police or the general terms and conditions of iHEARu-PLAY. To avoid abuse, a fair, understandable and open concept of data collection, storage, usage, and sharing was developed for iHEARu-PLAY, which will be described in the following.

3.1. Collected Data

From a user perspective, ‘privacy’ is a highly nuanced, culturally pre-determined and context-dependent social concept. An activity that is entirely acceptable and appropriate in one context might not be acceptable in another context. Eventually, the user’s own feelings and judgements have to be considered as guideline. Further, the idea that ‘the internet never forgets’ is extremely disturbing, given all the possible future uses of personal data. Therefore, it is ab-

solutely necessary to present an open and understandable concept of data collection to the user.

3.1.1. Data Types

Personally Identifiable and Mandatory Information:

Many companies assure their customers or their users of the service that collected personal data will be released only in a non-personally identifiable form. The underlying assumption is that ‘personally identifiable information’ is a fixed set of attributes such as names and contact information. iHEARu-PLAY goes one step further and will not ask the user for a name or address in the first place. Nevertheless, in order to use all functions of iHEARu-PLAY, users must register to the platform first. For this purpose, providing basic data is necessary, such as a freely chosen username, an associated e-mail address, and a password. Further, iHEARu-PLAY saves a user’s log-file for a duration of maximum seven days. This log-file will be deleted automatically after the given time and will just be used in cases of technical issues after being contacted by the user. Since the IP-Address of a user can easily be freely chosen and will not be stored for long time, the only traceable information extracted and stored could be the e-mail address, if the user’s actual name is encoded there.

Anonymous and Voluntary Information: In addition, further data might be disclosed voluntarily in the personal profile of iHEARu-PLAY, e. g., a user’s age, gender, personal health issues etc.; this is marked as optional information which – under certain circumstances – also might be used, for example for contacting the user, or through participation in surveys and feedback. iHEARu-PLAY stores the usernames and e-mail addresses in such a way that only selected employees, in detail researchers of the institute working on the iHEARu project² (Schuller et al., 2014b), have access to it. This assignment will only be used to identify a user’s data, if at a later time, the user likes his or her data to be deleted from the database. Data collection is always in accordance with applicable data protection regulations. The aggregated data will not be used to personally identify a user on the mentioned purposes.

Annotations: Annotations are collected using a smartphone application or a standard PC and can be done any time and anywhere as long as audio can be played back to the user. Even though iHEARu-PLAY’s primarily intended area of use is the labelling of audio databases, it is basically modality-independent, i. e., images and videos can also be imported.

Speech Recordings: Speech data is also collected using a smartphone application or a standard PC, which will allow the user to record prompted voice messages and upload them to the iHEARu-PLAY server – of course, only if the user’s explicit consent has been given. Microphones which are embedded in most laptop PCs, tablets, and smartphone devices can be used to perform the recordings. With this feature, collection of speech data under real-life conditions in the wild (e. g., different microphone types, devices, background noises, reverberations, etc.) of a large number of subjects with different geographic origins, languages, dialects, cultural backgrounds, age groups, etc. will be pos-

sible. Those speech signals collected in the wild may also contain different types of surrounding noises such as crowd noises from events, traffic noises, and other city noises.

3.1.2. Data Storage

In cases where personal information is entailed in some scenario, there should be complete guarantee that the delivered, stored, and transmitted data are managed only by the administrators with the appropriate access permissions. State-of-the-art technologies for secure storage in a locked server, delivery, and access of data will be used. Firewalls, network security, encryption, and authentication will be used to protect the collected data.

3.1.3. Data Access and Usage

All given voluntarily information, annotations, or speech data that a user creates are automatically saved by iHEARu-PLAY and internally connected to the user’s account. Therefore, all given data will be mapped to a pseudonym, which is internally also mapped to the username and e-mail address. All data will be stored electronically, always in an anonymised or pseudonymised form, and used exclusively for scientific research purposes; this means in particular:

- Access to a user’s username and e-mail address is restricted to the selected employees of the service.
- The information which maps the username and e-mail address to the related generated and given data will at no time be shared.
- The given user’s pseudonym and its linked pseudonymised data as well as the anonymised metadata will be shared with third party research bodies within and outside the EU only on a license base.
- Metadata, annotations, and recordings will be stored after the end of the iHEARu project (Schuller et al., 2014b) for use in follow up research projects. This will greatly help follow up research, ensure reproducibility of results, and eliminate the need to record new data over and over.
- Randomly selected samples of the pseudonymised audio data will be played to volunteers for annotation or for perception studies, either at a lab, or through crowdsourcing websites on the internet like iHEARu-PLAY.
- Samples of the audio data, generated figures, and anonymised metadata for statistical demonstrations can be used in scientific or public presentations. These figures can also be used in online scientific or public publications.

3.1.4. Data Changes or Deletion

The user has always the right to learn about all of his or her given information stored in the database, to correct it, or have them deleted. As a user, you can also access the relevant personal data at any time, change it, or remove it on your own. If the user disagrees with this general privacy policies, or wants to use the services of iHEARu-PLAY no

²<http://www.ihear.ru/>

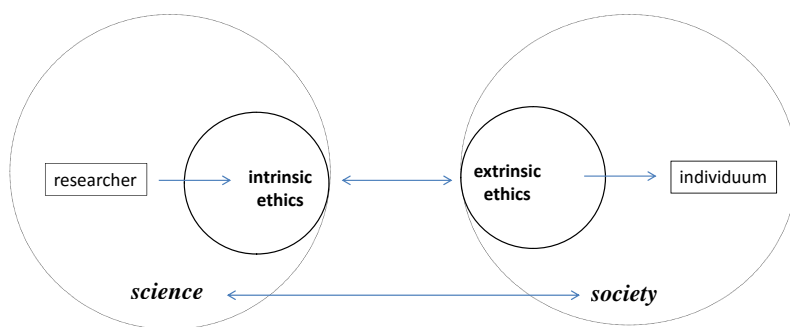


Figure 4: Intrinsic ethics (science) vs. extrinsic ethics (society).

longer, the user can request a deletion of the user account at any time. In the event of termination of the membership or a blocking of the account, iHEARu-PLAY will delete the e-mail address and the username of the user. Users always have the right to request the deletion of own annotations or recorded data through iHEARu-PLAY, and to obtain information about with whom the data have been shared. Nevertheless, data that have been already shared with third parties or that have been already used for publications cannot be deleted post hoc.

4. General Considerations

In this Section, we will broaden the view and present some general considerations on the role of science and an individual researcher (as part of science) on the one hand and society and individual (as part of society) on the other hand. In Figure 4, this interplay is depicted schematically. We tell apart intrinsic from extrinsic ethics, cf. (Batliner and Schuller, 2014): “In short, intrinsic ethics aims at producing sound scientific results; extrinsic ethics aims at the societal requirements that scientific results have to meet.”

Following the rules of intrinsic ethics or breaking them has both impact on society in general and – when an individual is directly involved – on the individual in particular. Intrinsic ethics pertains all aspects and criteria that have to be taken into account for producing ‘good’ science: That is what we learn in introductory courses, both at the beginner and at the post-graduate level, what we can read about in discussions in scientific journals; eventually, misconducts can lead to a public debate in newspapers and governmental bodies. Catchwords are: No plagiarism, sound reasoning (a somehow vague and generic, but very important requirement), adequate experimental design, adequate analysis and evaluation, correct use of (inferential) statistics (null hypothesis testing), adequate interpretation of results and taking into account possible impact, and last but definitely not least, adequate presentation of results to the public – for instance, by using ‘common language measures’ that can be conveyed easily to the non-expert (McGraw and Wong, 1992).

Possible impact leads to extrinsic ethics, where privacy considerations are in the fore for any study that employs individual subjects (or, in the case of big data, information that might be traced back to individual subjects); above,

we have detailed our approach within the iHEARu-PLAY game. Furthermore, which consequences it will have when we transfer results onto real life – for society in general and for some individuals in particular; think of the impact of a wrong modelling on therapy such as proof of concept of new drugs in humans with possibly detrimental consequences. Other examples of a direct impact on an individual is a wrong therapy based on faulty classification and subsequent modelling (recognition/generation/teaching of states such as emotions in the therapy of children with Autism Spectrum Condition, cf. Schuller et al. (2014a)), or sarcasm/emotion detection in conversations with automatic call-center agents. Besides such direct impact on the individual, there is indirect impact as well: misleading financing which prevents financing of promising approaches, and wrong societal decisions with unfavourable consequences for the individual.

Summing up, science (and every individual researcher) has an obligation to provide meaningful results – if it only were in exchange for the money given from society (public bodies, etc.). Now, we can as well turn the tables: “don’t ask what science can do for you – ask what you can do for science.” Provided that science really creates not only meaningful but also useful results, and especially in the case of important goals (for instance, diagnosis, and therapy of diseases such as speech pathology or autism), society should support science, and this means any individuals belonging to society as well. Naturally enough, this cannot be based on an obligation to deliver (such as taking part in experiments) but on the same terms as people are invited to donate blood – on a voluntary basis, with some incentives. This can be some payment for taking part in the experiments, credits for students, or – as we have illustrated above – simply fun in playing iHEARu-PLAY, while getting rewarded with prizes and scores and at the same time supporting science.

5. Conclusions and Outlook

We have shown ethic considerations concerning iHEARu-PLAY, a modular, cross platform, browser-based crowdsourcing game for collecting large-scale, real-life data for advanced paralinguistic tasks. When dealing with crowdsourcing, different ethical and privacy issues arise, e.g., concerning ethics in recording of private content, data col-

lection, annotation of crowdsourced data, and storage of the data. As the user will reveal personal data to iHEARu-PLAY, it is explained in detail what information and data iHEARu-PLAY collects, what usage the data has for us as researchers, with whom the data will be shared, and what will be done to protect a user's privacy. We also specified the measures taken to guarantee privacy in more detail. Moreover, we addressed general considerations on the role of science and researcher on the one hand (both have an obligation to provide results and data that have been financed by society), and society and individuals on the other hand (both should provide resources especially for important social issues). iHEARu-PLAY is still being developed; due to its modular architecture, rapid addition of new features is possible and planned. Besides integrating different kinds of features, our future work will focus on improving the quality management of the generated labels and recordings.

6. Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement No. 338164 (ERC Starting Grant iHEARu). We thank audEERING GmbH for technical support and sponsoring lottery prizes.

7. References

- Adda, G. and Mariani, J. (2013). Economic, Legal, and Ethical Analysis of Crowdsourcing for Speech Processing. In M. Eskénazi, et al., editors, *Crowdsourcing for Speech Processing: Applications to Data Collection, Transcription, and Assessment*, pages 303–334. Wiley, Chichester.
- Adda, G., Besacier, L., Couillault, A., Fort, K., Mariani, J., and De Mazancourt, H. (2014). "Where the data are coming from?" Ethics, crowdsourcing and traceability for Big Data in Human Language Technology. In *Crowdsourcing and human computation multidisciplinary workshop*, Paris, France, September. CNRS.
- Batliner, A. and Schuller, B. (2014). More Than Fifty Years of Speech Processing - The Rise of Computational Paralinguistics and Ethical Demands. In Cerna, editor, *Proceedings of ETHICOMP 2014*. 11 pages.
- Burkhardt, F., Eckert, M., Johannsen, W., and Stegmann, J. (2010). A database of age and gender annotated telephone speech. In Nicoletta Calzolari (Conference Chair), et al., editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).
- Eskenazi, M., Levow, G.-A., Meng, H., Parent, G., and Suendermann, D. (2013). *Crowdsourcing for Speech Processing: Applications to Data Collection, Transcription and Assessment*. Wiley, Chichester. 356 pages.
- Foundation, D. S. (2014 (Version 1.7)). Django web framework. Computer Software.
- Good, B. M. and Su, A. I. (2011). Games with a scientific purpose. *Genome Biology*, 12(12):135.
- Hantke, S., Eyben, F., Appel, T., and Schuller, B. (2015). iHEARu-PLAY: Introducing a game for crowdsourced data collection for affective computing. In *Proc. 1st International Workshop on Automatic Sentiment Analysis in the Wild (WASA 2015) held in conjunction with the 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 891–897, Xi'an, P. R. China, September. AAAC.
- McGraw, K. and Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, 111:361–365.
- Schmidt, F. A. (2013). The Good, the Bad and the Ugly: Why Crowdsourcing Needs Ethics. In *2013 IEEE Third International Conference on Cloud and Green Computing*, pages 531–535, Karlsruhe, Germany.
- Schuller, B. and Batliner, A. (2014). *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*. Wiley.
- Schuller, B., Marchi, E., Baron-Cohen, S., O'Reilly, H., Pigat, D., Robinson, P., Davies, I., Golan, O., Fridenson, S., Tal, S., Newman, S., Meir, N., Shillo, R., Camurri, A., Piana, S., Staglianò, A., Bölte, S., Lundqvist, D., Berggren, S., Baranger, A., and Sullings, N. (2014a). The state of play of ASC-Inclusion: An Integrated Internet-Based Environment for Social Inclusion of Children with Autism Spectrum Conditions. In Lucas Paletta, et al., editors, *Proceedings 2nd International Workshop on Digital Games for Empowerment and Inclusion (IDGEI 2014)*, Haifa, Israel, February. ACM, ACM. 8 pages, held in conjunction with the 19th International Conference on Intelligent User Interfaces (IUI 2014).
- Schuller, B., Zhang, Y., Eyben, F., and Weninger, F. (2014b). Intelligent systems' Holistic Evolving Analysis of Real-life Universal speaker characteristics. In Björn Schuller, et al., editors, *Proceedings of the 5th International Workshop on Emotion Social Signals, Sentiment & Linked Open Data (ES³LOD 2014), satellite of the 9th Language Resources and Evaluation Conference (LREC 2014)*, pages 14–20, Reykjavik, Iceland. ELRA.
- Seedhouse, D. (1998). Against medical ethics: a response to Cassell. *Journal of Medical Ethics*, 24(1):13–17.
- Silberman, M. S., Irani, L., and Ross, J. (2010). Ethics and tactics of professional crowdwork. *XRDS*, 17:39–43.
- Turk, A. M. (2016). Amazon mechanical turk participation agreement, 15.01.2016. URL: <https://www.mturk.com/mturk/conditionsofuse>.
- Zhai, H., Lingren, T., Deleger, L., Li, Q., Kaiser, M., Stoutenborough, L., and Solti, I. (2013). Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing. *J Med Internet Res*, 15(4):e73, Apr.