Proceedings of the workshop on Legal and Ethical Issues (Legal2020), pages 1–3
Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020

GDPR – A Game Changer for Acoustic Interaction analyzes

Ingo Siegert[1], Vered Silber Varod[2], Pawel Kamocki[3]
[1]Mobile Dialog Systems, Otto von Guericke University Magdeburg Germany
[2]Open Media and Information Lab, The Open University Israel, Israel
[3]Leibniz Institut für Deutsche Sprache, Mannheim, Germany
ingo.siegert@ovgu.de, vereds@openu.ac.il, kamocki@ids-mannheim.de

Abstract

Human interaction analyzes are essential to study social interaction, conversational rules, and affective signals. These analyzes are also used to improve models for human-machine interaction. Besides the pure acoustic signal and its transcripts, the use of contextual information is essential. Since the enforcement of the GPDR for the EU in 2018, there has been an increased uncertainty among scientists and participants. The discussion about the EU GDPR raised the awareness of personal rights and personal data recordings. This contribution aims to discuss issues of collecting personal and contextual data during acoustic interaction in terms of scientists' needs and GDPR demands.

Keywords: acoustic interaction, contextual data, GPDR, personal data

1. Introduction

The General Data Protection Regulation (EU Regulation 2016/679, hereinafter: the GDPR) entered into application on May 25, 2018. The aim of the GDPR is on one hand to unify all data protection laws across the European Union and on the other hand to protect the information about all EU residents against unlawful processing and privacy breaches. The GDPR has raised awareness about privacy-related rights throughout the EU. Data misuse has already led to noteworthy fines, including a 20k EUR fine against a French translation company Uniontrad for videotaping its employees, or against an Italian hospital (Azienda Ospedaliero Universitaria Integrata di Verona) of 30k EUR for not adequately protecting patient personal health records from unauthorized treatment. A danish taxi company stored data from eight million trips and thus violated the minimization principle of the GPDR, resulting in a fine of 161k EUR. The so-far highest fine (204M EUR) was condemned against British Airways due to a cyber incident, where 500,000 customers' personal data were compromised.

These examples of data misuse and the raising awareness have a direct impact on research activities, not only among scientists but also among experimental subjects.

The GDPR regulates the way data can be collected, stored, and processed (analyzed, exchanged, etc. ((Sveningsson Elm, 2009)). This entails the constitution of "personal data" and its efficient anonymization. According to Art. 4, 1. of the GDPR "personal data" is defined as "any information relating to an identified or identifiable natural person" ("data subject"). This concept of 'personal data' is significantly broader than the concept of 'personally identifiable information' (PII) used e.g. in the US. Personal data definition.

In contrast, interaction analyzes require huge data collections together with contextual information of the participants, in order to understand the interaction process, the individual behavior and develop proper models (Dudzik et al., 2019). These needs are challenging regarding the GDPR, leading to a huge uncertainty for data collection and data sharing activities.

Up until now, scientists tried to deal with it on their own and to help their peers by publishing documents and papers on ethical issues. Batliner and Schuller (2014), for example, list crucial ethical issues, including the challenge to guarantee the consent and the privacy of the subjects and the need to encode the data to guarantee this privacy (Batliner and Schuller, 2014).

This contribution aims to discuss the above issues by highlighting needs scientists have for analyzing interactions by giving examples in which additional "personal data" are needed but their storage and exchanging is crucial according to the GDPR.

2. Examples of the need for "personal data" in interaction analyzes

An important aspect of human perception is the processing of additional contextual information (Dudzik et al., 2019; Truong et al., 2007). The same holds true for technical systems. They must implement these human abilities and analyze human interaction signals together with additional contextual information. Therefore developers need databases capturing the context of interactions as well as the behavior expressed in them, which is also denoted as enriched data (Böck et al., 2019). Recent literature already surveys empirical research on how the decoding of behavioral signals in emotion perception benefits from contextual information (Wieser and Brosch, 2012) and how perceivers make use of contextual knowledge in interpreting affective behavioral signals (Aviezer et al., 2017). Important contextual categories are developed in (Dudzik et al., 2019), comprising age, gender, cultural embedding – nationality and ethnic background, language, and occupation. Furthermore, also personality traits, as NEO-FFI or SVF, and additional measurable signals are helpful, as it will be shown in the following.

Some examples where contextual data are needed to improve the interaction analyzes and modeling will be shortly discussed in the following. Age and gender information of a user is particularly required to improve automatic emotion recognition due to speaker-group dependent models (Siegert et al., 2014c) and could even improve multi-modal recognition for fragmentary data (Siegert et al., 2013). For example, in order to improve the emotion recognition significantly, recognition models make use of factors that affect the vocal tract, such as aging-effects, to improve their acoustic models by personalizing it to a specific age group. Also for human annotation of conversations not only the speech is important but also facial information (Siegert et al., 2014a), which allows an improvement of the identification of the participant.

Only a few analyzes so far deal with personality traits as additional contextual information. Although it is known that certain personality traits play an important role in communication (Cuperman and Ickes, 2009; Funder and Sneed, 1993; Weinberg, 1971). In (Gossen et al., 2017) the authors showed that the Incorporation of the contextual information on the personality trait "extraversion" improves the long term modeling of interactions. Furthermore, it is shown that information about the stress-coping ability of participants is useful to link exhaustive filled pause usage for the detection of challenging tasks (Siegert et al., 2014b). These examples underline the importance of both the acoustic signal and the contextual information (metadata) of the subjects to acoustic interaction research. Especially for automatic affect recognition systems the incorporation of metadata is beneficial.

3. GDPR-issues of recording contextual data

Recording contextual data of a participant can, even if all direct identifiers (name, birth, residence) are deleted, be used to identify a specific participant. A participant is identified when it's singled out from a group, typically by a sufficiently unique name-surname combination, but other identifiers (e.g., username or ID number, or in a certain context – a photograph) can also be taken into account. Moreover, a person is 'identifiable' if it can be singled out from a group by any means reasonably likely to be used (such as cross-referencing with data from social networks). Many examples are known, that not much data is generally needed to identify a person, even if the records are anonymized; e.g. the combination of zip codes, birth date and sex from anonymized data together with voter databases is enough to identify individuals (Ohm, 2010) Or that for identifying users of a famous video-streaming platform using knowledge about some movie ratings (Narayanan and Shmatikov, 2008).

What does that mean for the recording of contextual data? According to art. 5.1, c) of the GDPR, personal data should be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed'. The development of an experimental design and the conduction of the experiment can be very elaborate. Especially for interaction analyzes, as most of them are fundamental research, at the beginning of the experiment it is not clear which contextual factors are relevant for the task. Sometimes, additional questions arise during the evaluation of the data or new collaborative ideas are showing up during the presentation of the analyzes results. Thus, it is not always known nor desirable to limit the recording of personal data for research analyzes. This principle (which existed also under the 1995 Data Protection Directive), referred to as 'data minimization', is arguably the biggest hurdle for data-intensive research and technology especially as the GDPR does not allow any derogations from the principle for research purposes.

Furthermore, as the examples in the introduction showed, the protection of personal data against cyber incidents is crucial. Researchers mostly do not have the capacity nor the knowledge how to properly secure the data against misuse and on the same time still allow access to the data for authorized persons.

In this context, the German Research Foundation (DFG) for example explicitly encourages applicants to request funding for the preparation of research data for subsequent reuse or transfer. But this mostly covers the preparation, long-term archiving, and accessibility of data. Aspects relating to the compliance with the GDPR (access control, anonymization techniques, selective data access) are not in the focus so far.

4. Conclusion

The new regulations pose a new situation where despite not sharing personal data, or even not collecting at all, there is no proposed solution at the moment, at least for acoustic interaction research. Researchers might need a combined policy of legal and academic authorities.

One possibility is that the research community be more thorough, by disconnecting the assignment of context data to certain persons. This can be done by using ranges or broader classes for contextual data. For example, in (Silber-Varod et al., 2019) they used solely the acoustic signal and speaker-sex attribute, as the data was proprietary by an industrial company.

Another possibility is that recorded data is anonymized – hence the importance of anonymization or data omission for research activities gets important – as well as a proper access control infrastructure still allowing the share research data has to be developed. Hereby, it has to be noted that already the voice recordings itself reveal the speakers' identity. This constitutes a bigger challenge to cope with, especially in terms of anonymization.

Aviezer, H., Ensenberg, N., and Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current Opinion in Psychology*, 17:47 – 54. Emotion.

Batliner, A. and Schuller, B. (2014). More than fifty years of speech and language processing - the rise of computational paralinguistics and ethical demands. In *ETHICOMP 2014, 25-27 June 2014, Paris, France*.

Böck, R., Egorow, O., Höbel-Müller, J., Requardt, A. F., Siegert, I., and Wendemuth, A., (2019). *Anticipating the User: Acoustic Disposition Recognition in Intelligent Interactions*, pages 203–233. Springer International Publishing, Cham.

Cuperman, R. and Ickes, W. (2009). Big five predictors of behavior and perceptions in initial dyadic interactions:

Personality similarity helps extraverts and introverts, but hurts 'disagreeables'. *J Pers Soc Psychol*, 97:667–684.

Dudzik, B., Jansen, M., Burger, F., Kaptein, F., Broekens, J., Heylen, D. K. J., Hung, H., Neerincx, M. A., and Truong, K. P. (2019). Context in human emotion perception for automatic affect detection: A survey of audiovisual databases. In *Proc. of 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 206–212, Sep.

Funder, D. C. and Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *J Pers Soc Psychol*, 64:479–490.

Gossen, T., Siegert, I., Nürnberger, A., Hartmann, K., Kotzyba, M., and Wendemuth, A., (2017). *Modeling aspects in human-computer interaction - adaptivity, user characteristics and evaluation*, pages 57–78. Springer International Publishing, Cham.

Narayanan, A. and Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, May.

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review*, 57:1701–1777.

Siegert, I., Glodek, M., Panning, A., Krell, G., Schwenker, F., Al-Hamadi, A., and Wendemuth, A. (2013). Using speaker group dependent modelling to improve fusion of fragmentary classifier decisions. In *Proc. of 2013 IEEE CYBCONF*, pages 132–137, Lausanne, Switzerland.

Siegert, I., Böck, R., and Wendemuth, A. (2014a). Inter-Rater Reliability for Emotion Annotation in Human-Computer Interaction – Comparison and Methodological Improvements. *Journal of Multimodal User Interfaces*, 8:17–28.

Siegert, I., Haase, M., Prylipko, D., and Wendemuth, A. (2014b). Discourse particles and user characteristics in naturalistic human-computer interaction. In Masaaki Kurosu, editor, *Human-Computer Interaction. Advanced Interaction Modalities and Techniques*, volume 8511 of *LNCS*, pages 492–501. Springer, Berlin, Heidelberg, Germany.

Siegert, I., Philippou-Hübner, D., Hartmann, K., Böck, R., and Wendemuth, A. (2014c). Investigation of speaker group-dependent modelling for recognition of affective states from speech. *Cognitive Computation*, 6(4):892–913.

Silber-Varod, V., Lerner, A., Carmi, N., Amit, D., Guttel, Y., Orlob, C., and Allouche, O. (2019). Computational modelling of speech data integration to assess interactions in b2b sales calls. In *IEEE 5th International Conference on Big Data Intelligence and Computing (IEEE DataCom 2019)*, pages 125–127.

Sveningsson Elm, M., (2009). *How do various notions of privacy influence decisions in qualitative internet research?*, pages 69–87. SAGE Publications, Thousand Oaks.

Truong, K. P., van Leeuwen, D. A., and Neerincx, M. A. (2007). Unobtrusive multimodal emotion detection in adaptive interfaces: Speech and facial expressions. In Dylan D. Schmorrow et al., editors, *Foundations of Augmented Cognition*, pages 354–363, Berlin, Heidelberg. Springer Berlin Heidelberg.

Weinberg, G. M. (1971). *The psychology of computer programming*. Van Nostrand Reinhold, New York, USA.

Wieser, M. and Brosch, T. (2012). Faces in context: A review and systematization of contextual influences on affective face processing. *Frontiers in Psychology*, 3:471.