# Towards a New Methodology to Capture the Legal Compatibility of Conversational Speech Agents

Ernestine Dickhaut[†]
Information Systems
University of Kassel
Kassel, Germany
ernestine.dickhaut@uni-kassel.de

Laura Friederike Thies
Public Law
University of Kassel
Kassel, Germany
l.thies@uni-kassel.de

Andreas Janson
Information Systems
University of Kassel
Kassel, Germany
andreas.janson@uni-kassel.de

Alexander Roßnagel
Public Law
University of Kassel
Kassel, Germany
a.rossnagel@uni-kassel.de

Jan Marco Leimeister
Information Systems
University of Kassel
Kassel, Germany
leimeister@uni-kassel.de
Institute of Information Management
University of St.Gallen
janmarco.leimeister@unisg.ch

## ABSTRACT

Higher legal standards with regards to the data protection of individuals such as the EU General Data Protection Regulation (EU-GDPR) are increasing the pressure on developers of IT artifacts. This is especially prevalent when considering conversational speech agents (CSA), which are collecting data in new ways and thus are oftentimes producing conflicts with existing law regulations. For this purpose, we introduce the law simulation method, which is a well-known evaluation method among law researchers for capturing the legal compatibility of IT artifacts such as CSA. With this rigorous method, we are able to derive actionable guidance for CSA developers to evaluate developer efforts for increasing legal compatibility. To illustrate our methodological approach, we describe in this paper key steps of the method with respect to the evaluation of CSA. We briefly discuss how this procedure can serve as the foundation for a new evaluation method of legally compatible systems in information systems.

## CCS CONCEPTS

• **Human-centered computing** → **Interaction paradigms** → **Natural language interfaces** • Applied computing → Law

## KEYWORDS

Conversational speech agent, evaluation method, legal compatibility

## 1 INTRODUCTION

New IT artifacts bring new risks, which are sometimes not anticipated correctly because of the application situations or the behavior of the users [9]. In order to avoid them, the relevant social risks and consequences of a technology need to be identified and taken into account in the design process.

Typically, when developing systems, such as conversational agents, we subsequently evaluate them with users to elaborate on if our system design is appropriate, e.g., regarding guaranteed high usability, user experience perceptions, or outcomes of IT use (see especially [6] and [1] concerning our exemplary contributions related to the understanding, design, and evaluation of conversational agents). However, nowadays, other aspects are also important for system development. We are developing more and more systems that relate to legal compatibility and data policy issues. Consider, for example, the case of Amazon's Alexa, which activates itself when nobody is home or serves, as a consequence of its data collection efforts, as a witness in court (see also [11]). Nonetheless, we usually do not evaluate legal aspects when deploying these systems, oftentimes caused by the lack of appropriate evaluation methodologies for legal aspects when considering systems like CSA [5].

In this context, simulations are a great support for IT development. They help to visualize and play through abstract content quickly and without great effort [2]. As a rule, individual parameters can be easily varied to achieve the best possible results. In consequence, we draw on these advantages for

evaluating CSAs related to legal aspects. With the possibility to play through different system development parameters under realistic conditions, changes can be made relatively easily during the development process.

Thus, the method of simulation studies introduced by the law discipline [3] also provides a method-based foundation to evaluate technology such as CSA in a practical manner with regard to legal compatibility, thus making technology legally compatible. Therefore, we will propose in the following a comprehensive simulation methodology that provides developers of CSA with guidance when considering legal requirements, especially related to the EU-GDPR [4].

To illustrate this novel evaluation method, we developed a CSA as a smart personal assistant for university students with two overarching but somewhat conflicting design goals: (1) a high user experience and service quality that offers as much support during learning processes as possible but (2) also considers legal compatibility, i.e., achieving a higher legal standard than required by law. The prototype allows us to implement and evaluate our ideas for a (potentially) legally compatible speech agent. Our prototype can be classified, according to Knote et al. [8], into the category of an adaptive voice (vision) assistant. Our use case for deploying the CSA as well as the law simulation study method is a freshman course for economics and business administration that is taken by about 150 students. In the following, we will briefly describe the theoretical background we will draw upon, the law simulation study to evaluate legal compatibility for CSA, and a brief conclusion.

## 2 THEORETICAL BACKGROUND

To develop CSAs, we have to pay more and more attention to requirements from various disciplines, such as service quality and legal compatibility. One key aspect of speech agents is the usability, i.e., user experience, which we can sum up as the overall service quality. During the development process, these aspects are often well considered because, naturally, user-centered design considers those aspects in today's IT development methods. That is why service quality is frequently being paid much attention during the development process, while legal requirements are often addressed to a minimum extent in order to be compliant with the minimal requirements of law [7].

Legal compatibility goes further than mere legality and aims at the greatest possible compliance with higher-order goals to ensure that individual and collective social risks of technical systems are as low as possible. Consequently, it is to be distinguished from the term legality, which means the fulfillment of minimum legal requirements and is still common practice in many system development projects.

While legal compatibility aims to protect the user from risks, the quality of service is an essential factor for user satisfaction and thus also a decisive driver for the usability of a system. Quality of service describes the ability of a provider (or a system that replaces it) to meet customer expectations for a service delivery at a certain level of demand. The difference between expected and perceived factors is thereby an element of the continuous
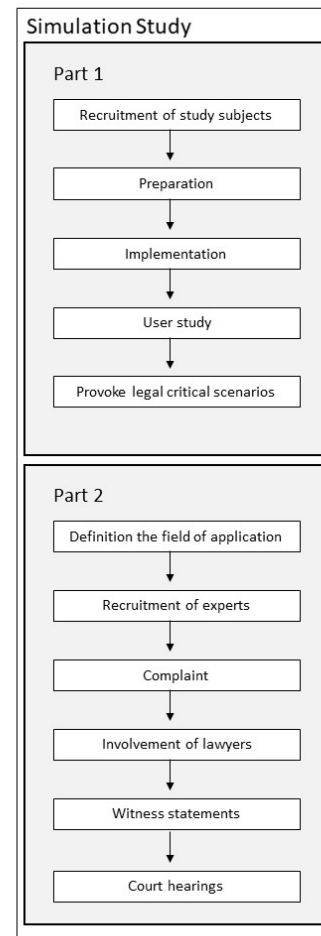


**Figure 1: Study Procedure**

spectrum of all possible requirement levels. For example, a user might prefer a CSA that is able to perform personalized services, such as playing their favorite song, based on their own preferences over a CSA that does not allow such personalization. Personalization of functionalities is a factor for service quality and thus an indicator for users of how satisfied they will be with the CSA during the use.

In times where data policy issues gain importance for providers of conversational agents, especially due to intensive fines in Europe when considering the EU-GDPR, there is a growing body of literature that recognizes the importance of the relevance of the consideration of data protection in technical systems [9], so we should keep legal aspects in mind during development. However, in practice, it is a challenge to consider it, especially when needing data for personalization aspects.

## 3 LAW SIMULATION STUDY

After developing a CSA, we typically evaluate our development efforts because that is the only way we can make a statement about the quality of the CSA (or generally IT artifacts). To evaluate the service quality, we know a lot of methods, like focus

groups, interviews, and questionnaires. For example, we can check the usability, user experience, and user satisfaction by conducting a user study with standardized questionnaires. To evaluate legal aspects for CSAs, we are faced with a new challenge because we do not know any validated method that helps us to sum up the implementation of the legal aspects. Therefore, we cannot make a statement about the extent to which we have achieved our goal to develop a legally compatible speech agent. The specific objective of this study was to find an evaluation method for legally compatible CSA. We use a law simulation study, which is a well-known method of evaluation among lawyers.

A law simulation study is a method to evaluate technology in a practical manner in regard to their legal compatibility. Similar forms of the law simulation study are already practiced in legal education and are called Moot Courts. In contrast to Moot Courts, the simulation study involves judges and lawyers who are already trained and have practical experience.

The procedure is characterized by the fact that it allows creating realistic usage situations while real damage is prevented [9]. This is achieved by letting test persons use prototype technology. The setting is to be designed as realistically as possible [9]. However, as there are no imminent dangers in the fictional usage situations for the test persons [10], it is desirable to provoke critical situations and situations of conflict that would not occur in such a high concentration [10]. In this way, simulation studies as a form of preventive technology design can make a substantial contribution towards sociotechnical design [3]. How the study is designed in detail highly depends on the technology that is being evaluated.

The law simulation study we perform, spanning the first and second quarter of 2020, serves the purpose of evaluating how our CSA works in real life in terms of legal compatibility. In this context, we want to find out how users and judges evaluate the technology we developed. The setup we used can be found in Figure 1.

Our law simulation study, therefore, consists of two parts and is partly based on [3]. The first part of the study is dedicated to how users evaluate the CSA regarding legal aspects. Moreover, this part of the law simulation study is suitable to gain knowledge about how the users perceive the service quality of the CSA. However, legal problems and concerns that only arise through practical use can also be identified in the first part of the study. In the second part of the law simulation study, we move slightly away from the users to get a reliable judgment on the legal compatibility of the CSA with the relevant laws. In this part, an oral court hearing will be conducted as realistically as possible with claims that could arise from the practical use of the CSA.

### 3.1 Part 1

The primary goal of the first part of the law simulation study is the use of the CSA by real users and capturing their perceptions while using the CSA. Our test persons, students in a basic course for economics and business administration, are recruited by sampling through self-activation. For the study, we have prepared

teaching material and have integrated it into our CSA. During the development, we paid special attention to legal aspects and service quality. In order to make the CSA as legally compliant as possible, special attention was paid to the protection of personal data. The data protection declaration has been presented in layman's terms and language to make it as understandable as possible. To provide the user with transparency about the data processing, the user will be informed at certain points in time about a) data that is stored, b) why this data is stored, and c) where this data is stored. Besides that, we interviewed the teacher of the course to get the best possible overview of his requirements for the speech agent. During the iterative development process, we also involved the teacher. The teaching material is prepared as a quiz to make it as comprehensible and supportive as possible.

In the first part of the law simulation study, we let students use our speech agent for half an hour a day for one week to revise the course material before the upcoming final assessment. We made the offer available for them voluntarily. Afterwards, we query the student's satisfaction with the assistant in terms of legal compatibility and service quality through a questionnaire. To prepare the second part of the study, we introduce legally critical situations like, for example, violations of the data protection declarations of the CSA, invalid data storages, or malfunctions. At the beginning of the first part of the study, students receive a short introduction to the CSA after which they can work with the CSA independently.

### 3.2 Part 2

In the second part of our law simulation study, we focus on the legal compatibility of the CSA we developed. For this purpose, we simulate a court trial that deals with and decides several questions concerning the legal compatibility of the CSA with relevant laws, in particular with the EU-GDPR. The violations from the user study lead to complaints on the part of the students in the further process of the study. To simulate the artificially created situation as realistically as possible, the plaintiff receives a lawyer. On the other side, the university as a defendant receives a lawyer as well. This is followed by a process that is usual for lawyers, in which various written pleadings are exchanged, the court is prepared, and the taking of evidence starts. Both sides prepare their statements and search for witnesses. The court will be staffed by law experts representing the judge and the attorneys on both sides.

The fictional cases the court is dealing with are built on the outcomes of the first part of our study. The situations of conflict we previously provoked will then be discussed and judged by the law experts, simulating a real court trial. We select a range of cases to trial during the law simulation study that are of high importance in the daily use of the technology. For instance, we could negotiate how voluntary consent can be given or how affected persons' rights can be enforced in a legally compatible way.

To make the simulation as realistic as possible, we invite four students who have already participated in the first part of the law simulation study. The students are supported by experts giving

legal advice while they take part as parties of the trial process. As a result of the second part of the law simulation study, we will get a judgment given by a law expert on several aspects of the CSA's legal compatibility.

## 3.3 Implications of the law simulation study

Based on the court's decision, we will derive an overview of the legal compatibility of our CSA and are able to say which possible risks arise from the design of our CSA. Overall, we expect that the results indicate improvement areas of our CSA and what aspects of system development need our special attention to ensure legal compatibility. Taking the two parts of the law simulation study together, we will derive a global statement on in how far our CSA is a) legally compatible and b) works in practice.

## 4 CONCLUSION AND FUTURE WORK

In conclusion, we briefly introduced the law simulation study and our exemplary implementation. On the one hand, we will expect that deploying this method for CSA developments enables us to receive feedback about the service quality and possible problems that occur during use. On the other hand, and more important because it offers us new possibilities in the development of legally compatible systems, the second part of the law simulation study will provide us with realistic feedback from a legal point of view. This feedback helps us to further improve our speech agent and avoid fines that are induced by stronger law regulations.

By conducting the study for the proliferating technology of CSAs, we will expect to provide broader findings about the applicability of the method with regards to the specifics of speech technologies. We will discuss and present those insights directly in the session to provide other HCI researchers interested in the development of speech technologies with actionable guidance for development. In this context, future work should also expand and test the method in law systems, since the method right now is adjusted to the European and German law system.

## REFERENCES

[1] Awad and Krishnan. 2006. The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. *MIS Quarterly* 30, 1, 13.

[2] Borges, G. 2018. Legal Framework for Autonomous Systems, 977–982.

[3] David, K., Geihs, K., Leimeister, J. M., Roßnagel, A., Schmidt, L., Stumme, G., and Wacker, A., Eds. 2014. *Socio-technical design of ubiquitous computing systems.* Springer, Cham.

[4] Dickhaut, E., Janson, A., and Leimeister, J. M. 2020. Codifying Interdisciplinary Design Knowledge through Patterns – The Case of Smart Personal Assistants. *Design Science Research in Information Systems and Technology (DESRIST).*

[5] Dickhaut, E., Janson, A., and Leimeister, J. M. 2020. Wie können Systeme künstlicher Intelligenz ohne Qualitätsverlust rechtsverträglich gestaltet werden? *WuM (Wirtschaftsinformatik & Management).*

[6] Hobert, S. and Wolff, R. M. von. 2019. *Say Hello to Your New Automated Tutor–A Structured Literature Review on Pedagogical Conversational Agents.* 14th International Conference on Wirtschaftsinformatik.

[7] Hoffmann, A., Schulz, T., Zirfas, J., Hoffmann, H., Roßnagel, A., and Leimeister, J. M. 2015. Legal Compatibility as a Characteristic of Sociotechnical Systems. *BISE* 57, 2, 103–113.

[8] Knote, R., Janson, A., Söllner, M., and Leimeister, J. M. 2019. Classifying Smart Personal Assistants: An Empirical Cluster Analysis. *Proceedings of the 52nd HICSS.*

[9] Pordesch, V., Roßnagel, A., and Schneider, M. 1999. Simulation Study Mobile and secure communication in Healthcare. *Datenschutz und Datensicherheit (DuD),* 76–80.

[10] Roßnagel, A. and Schuldt, M. 2013. The Simulation Study as a Method of Evaluating Socially Acceptable Technology Design, 108–116.

[11] Rüscher, D. 2018. Alexa, Siri and Google as digital spies on behalf of the investigation authorities?, 687–692.