An Educational Conversational Agent for GDPR

Behzad Mirzababaei[1(✉)] and Viktoria Pammer-Schindler[1,2]

[1] Know-Center GmbH, Graz, Austria
bmirzababaei@know-center.at
[2] Graz University of Technology, Graz, Austria
viktoria.pammer-schindler@tugraz.at

Abstract. Large-scale learning scenarios as well as the ongoing pandemic situation underline the importance of educational technology in order to support scalability and spatial as well as temporal flexibility in all kinds of learning and teaching settings. Educational conversational agents build on a long research tradition in intelligent tutoring systems and other adaptive learning technologies but build for interaction on the more recent interaction paradigm of conversational interaction. In this paper, we describe a tutorial conversational agent, called GDPRAgent, which teaches a lesson on the European General Data Protection Regulation (GDPR). This regulation governs how personal data must be treated in Europe. Instructionally, the agent's dialogue structure follows a basic GDPR curriculum and uses Bloom's revised taxonomy of learning objectives in order to teach GDPR topics. This overall design of the dialogue structure allows inserting more specific adaptive tutorial strategies. From a learner perspective, the learners experience a completely one-on-one tutorial session in which they receive relevant content (is "being taught") as well as experiences active learning parts such as doing quizzes or summarising content. Our prototype, therefore, illustrates a move away from the dichotomy between content and the activity of teaching/learning in educational technology.

Keywords: Educational conversational agent · Intelligent tutoring · General data protection regulation · Learning by argumentation · Bloom's taxonomy

1 Pedagogical and Technological Background

Lifelong learning is necessary for an individual, organisational and societal success and well-being. At the same time, increasing numbers of students in education or employees in workplaces in parallel to always seemingly too few resources make it challenging to provide a good level of individualised and interactive teaching. This, however, is desirable in order for teaching and learning to be of satisfactory quality [6,13]. Educational technology has long been investigated as a means to address this insight.

I. Hilliger et al. (Eds.): EC-TEL 2022, LNCS 13450, pp. 470–476, 2022.
https://doi.org/10.1007/978-3-031-16290-9_38

In this paper, we are particularly interested in the promise of conversational agents that act as tutors. Conversational agents constitute a human-computer interaction paradigm in which people can interact - so the ideal - in a relatively natural (for humans) way in natural language with technology. Ideally, with conversational agents, a learner can discuss concepts in a learning domain, move from talking about basics toward core complex definitions, do a self-assessment by answering questions, and receive feedback. This is what good educators do, given sufficient resources to interact bilaterally or with small groups of learners.

Much research in artificial intelligence for education has gone into developing computational systems that are able to, at least partially, fulfil some of these functions that (good) human tutors take on. Such systems are typically called intelligent tutoring systems [9,14]. More recently, researchers have investigated tutorial conversational agents, e.g., for question answering [7], helping students to efficiently use a large body of content [3], helping learners in assessing their own abilities [8], and providing administrative services such as answering students' questions on behalf of the academic faculty [10]. Many conversational agents that focus on teaching a topic are of course domain-specific, and by now research efforts span a plethora of subjects such as mathematics [2], medicine [11], computer science [16], physics and chemistry [17]. Typical research questions in these works are about the agents' architecture, how to model learners, different communication methods such as text or voice, or the impact of the appearance of agents on learners. Complementing such works, our research emphasis is on how to systematically design tutorial dialogues - which we propose to do by following Bloom's revised taxonomy of learning objectives [5] - and how to insert different teaching strategies into this overall structure.

This demo paper presents a conversational agent, named GDPRAgent, that carries out a complete tutorial conversation. The agent covers the complete content of a lesson step by step, asks questions after each step and gives feedback on learners' answers, and summarises content at the end of the lesson. GDPRAgent thereby simulates a whole learning session in a one-to-one situation between a teacher and a learner. The learning session is about the General Data Protection Regulation (GDPR), which is the European regulation that governs how personal data must be treated. The GDPR brings some new definitions and structures for data handling and management, as a result, individuals and organisations need to be adapted to the GDPR concepts. Therefore, it is a typical topic of MOOCs as it is relevant, at an introductory level, to a broad range of professions.

2 Description of the Prototype

GDPRAgent conveys the basic knowledge about the GDPR. The conversation contains four parts, which, topic-wise, we created to represent a typical introductory GDPR curriculum. First, it starts with a greeting and shows the agenda (Fig. 1). Second, the agent talks about what is the GDPR and where and when it should be applied. Third, the definition of personal data and sensitive data, and their differences are covered. Finally, the seven data protection principles of the

Fig. 1. The responsive web page in which the greeting section of the conversation and the agenda are shown.

Table 1. The flow of the conversation. It includes four subsections: greeting, the GDPR and its scope, personal and sensitive data, and seven data protection principles.

ID	Subsection
1.1	Greeting and agenda
2.1	What is the GDPR?
2.2	Where and when should it be applied?
3.1	What is personal data?
3.2	What is sensitive data?
3.3	What are the differences?
3.4	Asking relevant questions
3.5	Summarising information
4.1	Seven data protection principles
4.2	Asking relevant questions
4.3	Summarising information

GDPR are presented. Table 1 shows the order of the content in the whole dialogue. The GDPR content has been synthesised based on several online resources including the authoritative GDPR information[1]. Especially, we have benefited from FutureLearn[2], based on the open content license for non-commercial purposes, for the question parts of the dialogue, Parts 3.4 and 4.2.

Different teaching strategies are embedded into the conversation. In this prototype agent, we followed the revised Bloom's revised taxonomy of learning objectives [5]. For instance, in the third part of the conversation which is about personal and sensitive data (See Table 1), the agent first covers the "remem-

[1] https://eur-lex.europa.eu/homepage.html.
[2] https://www.futurelearn.com/courses/general-data-protection-regulation.

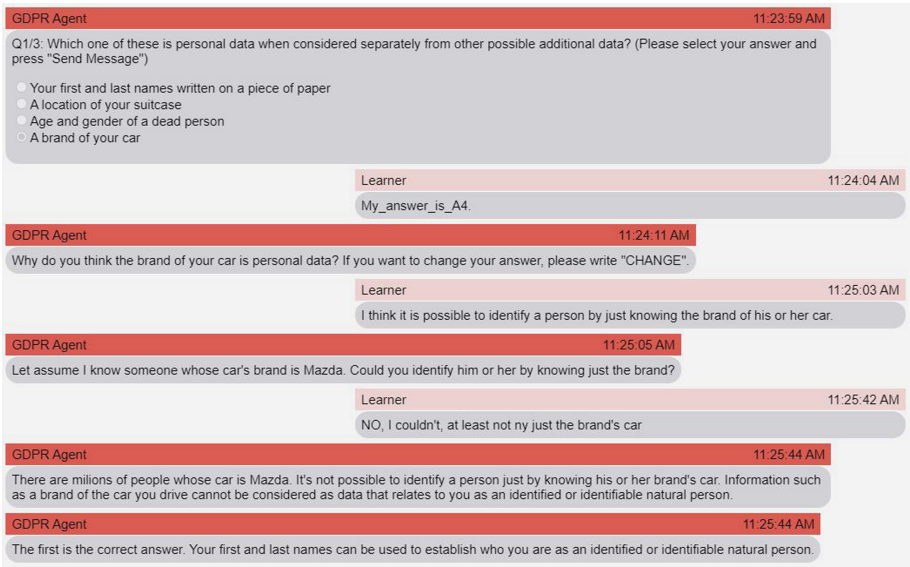

Fig. 2. The agent asks argumentative follow-up questions before showing the answer.

bering" and "understanding" levels of Bloom's taxonomy by talking about the definitions of personal and sensitive data. Second, by focusing on the differences between these two types of data and asking learners some questions in which the learners need to apply the information in various scenarios, the agent covers "applying" and "analysing". Finally, the agent addresses the "evaluating" level by having an argumentative conversation in which the learners are asked to justify their answers. At this point, we therefore also followed the teaching strategy of learning through argumentation [12]. Learning through argumentation guides learners to analyse a problem from various perspectives and also to distinguish what is correct and incorrect. Figure 2 shows how the agent asked follow-up questions in order to guide the learner to find out why the selected option was incorrect. Here the agent asked the user to justify his answer and then the agent explained a situation in which the user's answer is not valid. In general, the agent adapts to learners' responses. Based on each response, the agent asks the learner to think again about their own response and justify it and then, in case of selecting a wrong answer, the agent explains a situation in which the user's argument is not valid anymore. This part of the dialogue uses learning by argumentation.

The example of the "learning through argumentation" teaching strategy above already shows, how an overall instructional design following Bloom's revised taxonomy of learning objectives allows and actually needs the insertion of more specific teaching strategies. Note that from an instructional perspective "teaching strategies" are inserted, whilst this means that from a technical perspective "adaptation mechanisms" need to be inserted. Here, the full spectrum of intelligent tutoring and adaptive teaching systems [4] is available to conversational agent designers.

Fig. 3. The agent asks for the learner's idea about the GDPR's scope before teaching it. The agent is also adapted to the learner's responses

In the current prototype, we have inserted two more adaptation strategies. The first is to adapt to learner knowledge (cp. [4]' taxonomy of adaptation strategies), and to exercise what a learner does not know. We did this in Parts 3.4 and 4,2 (See Table 1), such that when the agent asks questions about the different type of data (Part 3.4) and about data protection principles (Part 4.2), the number of questions for each learner depends on the number of his or her correct answers. We defined five different questions for Part 3.4 and six questions for Part 4.2, but the agent first asks three questions. In each part, if a learner answers at least two questions, the agent asks the learner for answering more questions. In case of agreement, the agent asks the rest of the questions. Otherwise, the conversation is continued. The second adaptation strategy could be understood as an adaptation that targets learners' affect. At the beginning of a new topic, the agent asks what the learner already knows about this. In Fig. 3 for instance, the agent asks the learner, about the GDPR's scope before giving the information. The agent is to some extent adaptive to the learner's responses. The agent uses keyword matching in order to understand the learners' responses. For each topic, a set of keywords are defined which helps the agent to have an adaptive reaction. For instance, in Fig. 3, since the user did not know the answer, the agent gave encouraging feedback in order to motivate the learner. In general, if the agent does not understand the user's responses, it will try to keep the conversation coherent and meaningful by giving a proper reply.

Technically, we have implemented GDPRAgent based on the open-source Bazaar framework [1] as back-end[3], and as an HTML/JS responsive web page for the front-end. This framework allows both rule-based and machine-learning-based classifiers to decide between dialogue branches. GDPRAgent is ready to use and publicity available[4].

[3] https://github.com/DANCEcollaborative/.
[4] http://chatbot.know-center.tugraz.at/bazaar/landing_page/chatbot_landing_page.

3 Future Work and Vision

In ongoing work, we are working on assessing the agent's usability, which is an important baseline that educational technology needs to meet. We are further working on investigating what qualities of the learning process and learning outcomes change as the interaction paradigm is more conversational when compared to other interactive digital content formats.

We see the main contribution of our research to existing research on conversational agents in education in the systematic instructional dialogue design, based on Bloom's revised taxonomy of learning objectives. We aim to show that this structure can also be used in other domains than the GDPR. Further, above we have explained already a few teaching (instructional perspective) and adaptation (technical perspective) strategies. A systematic guideline for educational conversational agent developers would be helpful that summarises which teaching strategies can be inserted in a single conversational agent lesson. Finally, we have been working on natural language processing capabilities that allow processing and feedbacking more open-ended questions of a particular argumentative form [15] and think that there is overall still room for improvement in research on being able to accommodate more complex question types and feedbacking them in intelligent tutoring systems.

Overall, we see the promise of such fully conversational intelligent tutoring systems as moving beyond the content/learning dichotomy, that separates the provision of content with the support for active learning activities in computational environments for learning. GDPRAgent can demonstrate what such an educational technology could look like in the future. As a note of caution: We are thereby not supporting the stance that human teachers can or should be replaced. Readers will note that our agent teaches the basics of GDPR. Given the instructional and content design effort that goes into creating a conversational agent such as ours, which ultimately covers just a single tutorial unit, we foresee that such agents will rather replace, or become the norm in, learning content management systems and MOOCs, which already step ahead of very traditional content-focused computational environments for learning.

Acknowledgements. This work was supported by the "DDAI" COMET Module within the COMET-Competence Centers for Excellent Technologies Program, funded by the Austrian Federal Ministry (BMK and BMDW), the Austrian Research Promotion Agency (FFG), the province of Styria (SFG) and partners from industry and academia. The COMET Program is managed by FFG.

References

1. Adamson, D., Dyke, G., Jang, H., Rosé, C.P.: Towards an agile approach to adapting dynamic collaboration support to student needs. Int. J. Artif. Intell. Educ. **24**(1), 92–124 (2014). https://doi.org/10.1007/s40593-013-0012-6
2. Aguiar, E.V.B., Tarouco, L.M.R., Reategui, E.: Supporting problem-solving in Mathematics with a conversational agent capable of representing gifted students'

knowledge. In: 2014 47th Hawaii International Conference on System Sciences, pp. 130–137. IEEE (2014)
3. Akcora, D.E., et al.: Conversational support for education. In: AIED 2018. LNCS (LNAI), vol. 10948, pp. 14–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93846-2_3
4. Aleven, V., McLaughlin, E. A., Glenn, R. A., Koedinger, K.R.: Instruction based on adaptive learning technologies. Technical report (2017)
5. Anderson, L.W., Krathwohl, D.R., Bloom, B.S.: A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives, p. 352 (2001)
6. Brinton, C.G., Rill, R., Ha, S., Chiang, M., Smith, R., Ju, W.: Individualization for education at scale: MIIC design and preliminary evaluation. IEEE Trans. Learn. Technol. **8**(1), 136–148 (2014)
7. Clarizia, F., Colace, F., Lombardi, M., Pascale, F., Santaniello, D.: Chatbot: an education support system for student. In: Castiglione, A., Pop, F., Ficco, M., Palmieri, F. (eds.) CSS 2018. LNCS, vol. 11161, pp. 291–302. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01689-0_23
8. Durall, E., Kapros, E.: Co-design for a competency self-assessment chatbot and survey in science education. In: Zaphiris, P., Ioannou, A. (eds.) HCII 2020. LNCS, vol. 12206, pp. 13–24. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50506-6_2
9. Gertner, A.S., VanLehn, K.: Andes: a coached problem solving environment for physics. In: Gauthier, G., Frasson, C., VanLehn, K. (eds.) ITS 2000. LNCS, vol. 1839, pp. 133–142. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-45108-0_17
10. Hien, H.T., Cuong, P.N., Nam, L.N.H., Nhung, H.L.T.K., Thang, L.D.: Intelligent assistants in higher-education environments: the fit-ebot, a chatbot for administrative and learning support. In: Proceedings of the Ninth International Symposium on Information and Communication Technology, pp. 69–76 (2018)
11. Jagtap, R., Phulare, K., Kurhade, M., Gawande, K.S.: Healthcare conversational chatbot for medical diagnosis. In: Handbook of Research on Engineering, Business, and Healthcare Applications of Data Science and Analytics, pp. 401–415. IGI Global (2021)
12. Jonassen, D.H., Kim, B.: Arguing to learn and learning to argue: design justifications and guidelines. Educ. Tech. Research Dev. **58**(4), 439–457 (2010). https://doi.org/10.1007/s11423-009-9143-8
13. Kim, B.: Social constructivism. Emerg. Perspect. Learn. Teach. Technol. **1**(1), 16 (2001)
14. Koedinger, K.R., Anderson, J.R., Hadley, W.H., Mark, M.A.: Intelligent tutoring goes to school in the big city (1997)
15. Mirzababaei, B., Pammer-Schindler, V.: Developing a conversational agent's capability to identify structural wrongness in arguments based on Toulmin's model of arguments. Front. Artif. Intell. **4** (2021)
16. Mitrovic, A.: An intelligent SQL tutor on the web. Int. J. Artif. Intell. Educ. **13**(2–4), 173–197 (2003)
17. Pérez-Marín, D., Boza, A.: A procedure to create a pedagogic conversational agent in secondary physics and chemistry education. Int. J. Inf. Commun. Technol. Educ. **9**(4), 94–112 (2013)