Tutoring Postgraduate Students with an AI-Based Chatbot

https://doi.org/10.3991/ijac.v16i1.35437

Matti Koivisto(✉)
South-Eastern Finland University of Applied Science, Mikkeli, Finland
matti.koivisto@xamk.fi

Abstract—Tutoring and counselling are essential support services that universities provide to their students. With the ever-increasing number of the students, providing these functions has become harder and harder. Therefore, universities have started to use chat robots, or chatbots, to automate their guidance and counselling services. This paper discusses the suitability of the chatbots or virtual advisers in personal study planning and course selection in higher educational institutions. The empiric part of the study reports the observations of the experiment conducted in a Finnish university of applied sciences. In the study, postgraduate students of engineering used a tailor-made chatbot while creating their personal study plans. The findings of the study indicated that chatbots could, to some extent, improve student counselling, and the main advantages were scalability and unlimited service hours. However, students did not see that artificial intelligence could fully remove the need for human counselling. The main reported shortcomings of the chatbot were the minor significance of the individuality, lack of inspiring effect, and general negative attitude towards automated services.

Keywords—chatbot, tutoring, counselling

1 Introduction

Chatbots are, today, more popular than ever. They have a growing presence in modern society, and we can find them in many tasks for example in product and service industries, health care, and medicine as well as education. Two main reasons for the increased popularity of the chatbots are advances in artificial intelligence (AI) and the growth of different kinds of mobile text-based messaging applications [1]. Earlier studies have indicated that universities have used the AI mainly in the following four application areas: adaptive systems and personalization, assessment and evaluation, profiling and prediction, and intelligent tutoring systems [2]. This paper focuses on the use of chatbots in tutoring and counselling services, and more precisely, it analyzes the suitability of virtual advisers in personal study plan creation and course selection in HEIs.

Universities and other educational institutions provide a large variety of tutoring services to their students. Traditionally, the members of the staff have had an essential role in providing these services, but over the past few decades, the work has shifted more to dedicated specialists [3]. Tutoring provided by trained people has many benefits, but it is very labor intensive, and tutors can provide their services only during limited service hours. With the ever-increasing number of students, universities have sought to streamline their own operations with automation. Earlier studies have provided mixed results on chatbots in education [4]. Therefore, more studies are needed to investigate the possibilities of virtual assistants in HEIs.

To facilitate insights into this topic we organized our paper as follows. Section 2 of the paper introduces the principles of the chatbots and some earlier studies on using chatbots in higher education. Since practically all students must select courses to determine their study program, our aim is that the virtual assistant is accepted and used on a large scale. Thus, in Section 3, we briefly cover the main theories on technology acceptance and the metrics created by scholars and practitioners to predict the success of new services. Our empirical study is presented in Section 4. In our experiment, we first develop a student counselling chatbot called Vivian and then analyze postgraduate students' opinions on its suitability to provide advice on course selection and curriculum planning. Finally, the work ends with the conclusions of Section 5.

2 Artificial intelligence and chatbots

In recent years, AI has become a real buzzword that has gained more and more public attention. But what do we mean by AI? Although scholars have suggested many definitions, there is no widely accepted definition of AI [5]. The challenge has not been in how to interpret the first letter "A" but the trouble comes mostly from the "I" [6]. However, intelligence is often seen as the ability to learn and to solve problems in an uncertain context [7]. In AI, the scope of learning can be limited to some specific area (narrow AI) or it can aim to achieve complex goals in a wide range of environments [8]. In this latter case, we talk about general AI, which aims to act and think like a human in a wide range of cognitive tasks.

2.1 Chatbots and human-AI collaborative conversation

In this paper, we focus on one narrow AI application, namely chatbots. A chatbot is a computer program designed to simulate the discussion with humans. Chatbots can be divided into two categories: rule-based and AI chatbots. The rule-based chatbots use keywords in language understanding, and their programming is based on simple if-then statements. The AI-based chatbots, instead, apply natural language processing and multi-level intent hierarchy [9]. The main advantage of AI chatbots is their scalable learning capabilities based on machine learning and therefore during the recent years the focus has entirely shifted to AI based chatbots.

Chatbots can be used for many purposes, but they can be classified to two main categories: question-answering assistants and social bots [10]. The question-answering bots are task oriented, and they provide answers and solve specific problems based on their information [11]. Many organizations use them in different kinds of customer service tasks, and according to some estimates, in 2020, already 25% of the buyers used chatbots to communicate with businesses [12]. Social chatbots are instead intelligent dialogue systems that can engage in empathetic conversations with humans [13], and these systems are supposed to provide personalized responses for the users in their areas of interests. Some popular examples of social chatbots are Woebot Health, a personal mental health ally [14], and XiaoIce, a Chinese social chatbot with more than 660 million users [15].

Chatbots communicate with humans using natural language either with a text or voice interface [16]. Although many organizations have published different kinds of guides and rules how to design a successful bot, there are still many open questions in chatbot design [17]. The main challenges are related to grammatical complexity and semantics variations of human communication (for more details see ref. [18]). However, earlier studies suggests that successful chatbots must meet the following requirements [11]:

- They must have conversational capabilities beyond short answers.
- They must offer semantically correct information and use context specific vocabulary.
- They must give meaningful responses and answers.
- They must be trained with a domain-specific dataset.

These requirements highlight that chatbot designers must pay attention not only to functionality but also to the social aspects [19] and to the form features of the chat-bots [20]. These features together affect user experience and play an essential role in the user satisfaction and service acceptance.

2.2 Chatbots in education

Over the past few years, universities have used chatbots both in formal and informal learning and in many other functions. Students have used educational chatbots, for example, in language, engineering, science, nursing, social science, and business studies [21]. In addition, higher education institutions (HEIs) have applied chatbots for example library services [22], student counselling [23], admission, and career guidance [24], and they apply chatbots repeatedly in new areas. Figure 1 shows Science Direct and Ebsco database search result for two different search criteria (chatbot and chatbot & education), and it clearly points out that the interest towards chatbots, both in general and in education, has clearly increased during the last few years.

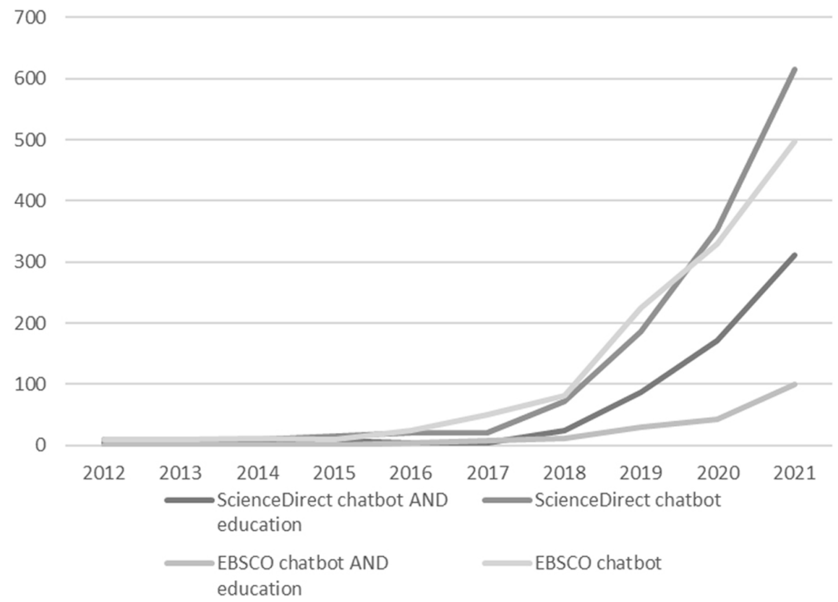

Fig. 1. Development of chatbot related publications during the last 10 years

Earlier studies have identified both the benefits and limitations of chatbots in HEIs. Some scholars have found that chatbots increase students' learning motivation [25] and attention [26] and that they support peer communication [27] and collaborative learning [28]. However, some researchers have reported the limitations of the chatbots. For example, current implementations seldom provide enough personalization [29], and users seem to trust chatbots less in more complicated tasks [30].

In this paper, we focus on the use of chatbots in personal curriculum planning and course selection. Course selection is a sequential decision-making process in which students determine their study programs [31]. During the selection process, students typically need some form of guidance. Traditionally, this support has been provided either by faculty members or by designated counsellors. Although human counselling has many benefits, it has its limitations. Earlier studies have identified for following challenges in traditional advising: recommendations can be inconsistent across counsellors [32], it is time consuming [33], it is not available outside office hours [34], and the advising sessions are limited in time, especially during the highest-demand periods [33].

3 Acceptance of information systems

There are two main reasons why HEIs have started to apply chatbots for tutoring and counselling services. Firstly, they want to offer better services to their students. Secondly, they want to use their limited resources as efficiently as possible. For universities to achieve these goals, students must adopt the new services on a large scale. However, which factors make it possible for users to accept the information system?

3.1 Different approaches to study information system acceptance

Scholars have spent a lot of time and effort to study the factors and processes affecting users' adoption and use of information technology (IT). They have looked at the issue mainly from one of the following three perspective: technology acceptance, information system success, or usability. Although all these methods seek to understand why users accept or reject certain services and how user acceptance is affected by system design features, they use different methods to achieve this goal [35].

IT acceptance studies have focused on different sets of acceptance determinants, and scholars have created multiple models like Innovation Diffusion Theory [36], Technology Acceptance Model [37], and Theory of Reasoned Action [38]. Although scholars have identified different acceptance determinants, the basic concept underlying them links individual reactions and intentions to the actual use of the system [35]. In organizational settings, the acceptance of the new services does not occur in isolation, but social influence plays also an important role. It is conventional wisdom that friends and peers have a large influence on us. This holds true especially when an individual's underlying behavior is social or networked [39].

Information system success studies instead focus on the impacts of the system either at an organization or individual level. At the organizational level, the success has been measured, for example, with economic metrics like return on investment or pay-back time. At the individual level, researchers have relied mostly on user satisfaction, as an indicator of IS success [35].

Although system acceptability is a large multidimensional phenomenon [40], Human-Computer Interaction (HCI) professionals have analyzed it mainly from the usability perspective. The most important difference between information system success and HCI studies is the time factor. IS scientists are typically operating with implemented products or services whereas usability professionals address the process leading up to the implementation of the system. This process, often referred to as usability engineering, includes all steps required to produce usable products [41].

3.2 Metrics of our study

Researchers from different disciplines have developed a range of feedback metrics and methods for predicting service acceptance. A typical study enumerates the attributes of the system, collects data from users, and analyzes them separately, making it a potentially useful diagnostic for system design. However, earlier studies have indicated that any single metric like user satisfaction is a weak predictor of system use [42], but IS success measurements should combine multiple dimensions like satisfaction, system use, system quality, and benefit constructs.

In our study, we will use three different metrics to analyze the relationship between the service and its users. These metrics are Customer Satisfaction (CSAT), Customer Effort Score (CES), and Net Promoter Score (NPS), which are currently among the most popular techniques for measuring the strength of the relationship between the company's services and its customers [43].

CSAT uses a five-point Likert scale to measures the short-term user satisfaction. It targets a "here and now" reaction; therefore, it has its limitations. Scholars have found out that it has limited capability to measure user's long-term relationship with the service or service loyalty [44]. CES instead measures how much or how little effort is needed from the user to use the service and it uses Likert scale ratings from 1 to 7 [44]. It is important to understand that CES analyzes the user experience from a different perspective than CSAT. Instead of focusing on maximizing the satisfaction, it pays attention to minimizing the effort [43]. Our third metric, NPS, measures long-term loyalty and determines which customers are likely to promote our service to others; therefore, it is often used as a growth indicator. NPS has a scale from 0 to 10, and it groups the responders to the following four categories: "promoters" (9–10 rating – extremely likely to recommend), "passively satisfied" (7–8 rating), and "detractors" (–6 rating – extremely unlikely to recommend) [45]. Findings of the NPS surveys are typically reported with a NPS score which is calculated according to Equation 1.

$$NPS\ score = \frac{(Number\ of\ promoters - Number\ of\ detractors)}{Number\ of\ responses} \quad (1)$$

4 Empirical part of the study

4.1 Introducing Vivian

The aim of our study was to analyze the suitability of the tutoring chatbots in higher education. To be able to do this, we first developed a simple study counselling chatbot Vivian. Vivian is a typical question-answering assistant, and it offers its suggestions and advice on questions related to personal curriculum planning and course selection questions. Vivian was implemented using IBM's Watson Assistant cloud service. Watson Assistant uses AI and "has the ability to answer commonly asked questions, based on domain and industry specific information" [46].

The basic structure of our chatbot is shown in Figure 2. Users interact with Vivian using their web browsers either on their mobile phones or laptops. All communication between the user and Vivian is text based. When the chatbot receives a message from the user, a logical structure called a dialog skill interprets the user input and directs the flow of the conversation. The dialog skill consists of three components, which are intents, entities, and dialogs. As the name implies, the intent component tries to recognize the intent of the user or in other words what the user wants to know. Vivian detects for example the following intents:

- #Structure (questions related to the structure of the curriculum),
- #Mandatory (questions related to the compulsory courses),
- #Optional (questions related to the optional courses),
- #Thesis (details of the master's thesis)
- #Recommendation (questions related to the course recommendations using different criteria), and
- #Skills (suggestions how to learn a specific new skill).

When an intent is created, the designer provides the initial configuration. For example, #Thesis intent is initially identified with the following kinds of questions:

- Can you provide me more details on the final work?
- Could you tell me more about final thesis?
- How large is the thesis?
- Can I have more information about final work?
- What are the requirements of the diploma work?
- What is master's thesis?

But because our chatbot is based on AI, it learns new ways to detect the user's intent. It can, for example, learn that a user input *"I want more info about final work"* means that the user's intention is to learn more about master's thesis.

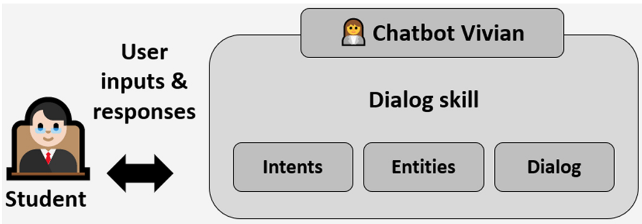

Fig. 2. The basic structure of the chatbot

Entities that can also be called keywords are used to catch essential information from the user input. Vivian recognizes for example the following entities:

- @Study_type, which refers to type of a course (possible values: compulsory, optional, thesis, etc.)
- @Interest, which specifies the interest of the student (possible values: leadership, entrepreneurship, research methods, decision-making, digitalization, etc.).
- @Skill, which identifies the new skill the student wants to learn (possible values: teamwork, academic writing, environmental, etc.).

Each value has many synonyms that are either added by the user or recommended by the Watson Assistant. For example, one value of the skill entity called *Environmental* has synonyms like green, ecological, sustainability, environment, and climate.

Finally, the third components of the dialog skill are called dialogs, and they are nodes or action blocks in a dialog tree. All dialogs created by the chatbot designer are placed between two automatically created nodes *Welcome* and *Anything else*. *Welcome* node is the starting point of all conversations, and it contains some form of greeting. Vivian, for example, had seven different and randomly used ways to introduce itself in the beginning of the communication. *Anything else* node is the last resort, and it is used if Vivian does not understand the user input. If our chatbot fails to interpret the user's intent, it will first ask the user to rephrase the question. Because Vivian tries to mimic a human assistant, it does not always use the same phrases but varies its messages. If Vivian is not able to understand the user after the second time, it recommends user to rephrase the question or to send an email to his or her tutor.

The relationship between the dialog skill and the messages sent to the user is demonstrated in Figure 3. At the beginning of the conversation, Vivian first introduces herself with Welcome node and then asks the name of the user with Get name dialog.

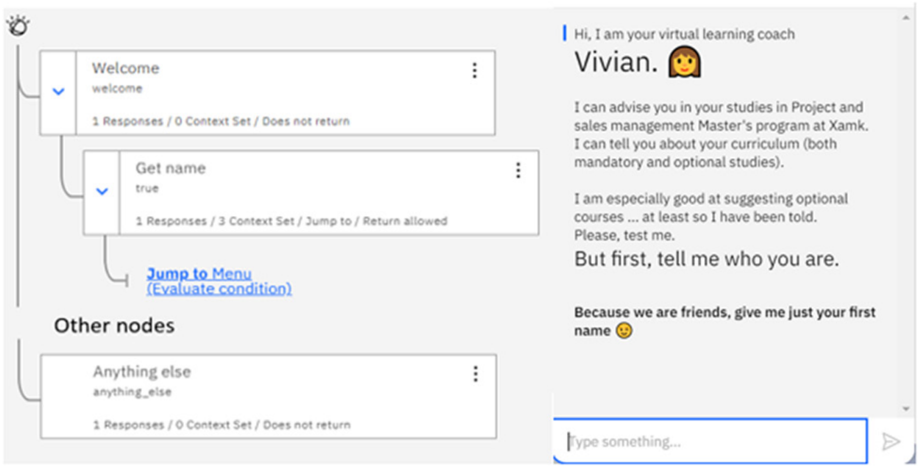

Fig. 3. Welcome and Get name dialogs or nodes in the dialog skill and their outcome

In typical conversations, students first asked Vivian to show the structure of the curriculum (#Structure intent), and then they quickly looked at some mandatory courses (#Mandatory intent). However, Vivian's main task was to advise students to choose optional courses, and therefore, students spent most of their time using this feature. Students were able to ask Vivian to list all available optional courses (#Optional intent), to give recommendations based on their own interest (#Recommendation intent), or willingness to learn new skills (#Skill intent). During this process, they could use some pre-defined categories or open questions or statements. Vivian analyzed the request, and if it was able to find a suitable course or courses, it responded with best matches. If students did not specify any area of interest or new skill to learn, Vivian used its Surprise me functionality and suggested a random course. If Vivian did not find any match for the student request, it apologized and offered three options for moving on. These alternatives were sending a new question, reviewing the courses offered by other universities of applied sciences, or sending an email to a human tutor.

4.2 Study design

In the empirical part of our study, we carried out an experiment at a Finnish university of applied sciences to analyze how well Vivian can help our students in their curriculum planning tasks. In our study, postgraduate engineering students used the chatbot when they planned their studies and selected optional studies from a relatively large pool of alternatives. We introduced our chatbot to our students during a normal online learning session. After that, they were asked to use the service if they wanted and then fill the feedback questionnaire. In total, 57 students took part on the online session, and 53 of them decided to use Vivian and filled the feedback form. Our questionnaire had three parts. In the first part, we asked about the user experience and used the earlier introduced metrics CSAT, CES, and NPS. The second part concentrated on chatbot specific topics, and the questions were based on the requirements of the natural and successful dialogue discussed in Section 2.1. In the third part, students were asked to evaluate the strengths and weaknesses of the chatbots in the course selection task.

4.3 Results from the students' perspective

The results of the first part of the questionnaire are shown in Table 1. Because all three metrics used different scales, we present the results both with original values and normalized to the scale 0 to 1. Our data revealed that the students found Vivian quite easy to use (CES = 0.80), but the satisfaction score was clearly lower (CSAT = 0.68) and students were unlikely to recommend it to other users (NPS = 0.63). The differences between CES and other two metrics was statistically significant ($p = 0.05$). The findings of the NPS surveys are not typically reported using a mean but the NPS score is used instead. Therefore, we also calculated the NPS score. The NPS score of our chatbot was remarkably low –0.38, or –38%, because only three students were ready to promote it and 23 students in total could be classified as detractors.

Table 1. Findings of the general metrics

Metric	Mean Original Scale	Mean Scale 0 to 1	SD Scale 0 to 1
Customer Effort Score (CES)	5.60	0.80	0.16
Customer Satisfaction (CSAT)	3.40	0.68	0.12
Net Promoter Score (NPS)	6.26	0.63	0.19

With the second part of our questionnaire, we collected data on students' opinions on chatbot specific topics introduced earlier. The results are summarized in Table 2. According to the results, students gave the lowest score to Vivian's conversational capabilities. The other dimensions had quite similar ratings (in the range from 0.73 to 0.76). The difference between conversational capabilities and other metrics was statistically significant ($p = 0.05$).

Table 2. Findings of the chatbot specific metrics

Metric	Mean Original Scale	Mean Scale 0 to 1	SD Scale 0 to 1
Conversational capabilities	3.11	0.62	0.16
Correct information and vocabulary	3.66	0.73	0.12
Meaningful answers	3.74	0.75	0.14
Case specific information	3.87	0.76	0.20

In the third part of the survey, students identified the strengths and weaknesses of the chatbot in course selection task with open-ended questions. The four most-often-mentioned Vivian's strengths and weaknesses are listed in Table 3.

The frequency describes how many students from all participants mentioned this topic. It is worth pointing out that Vivian's ability to work 24/7 was an extremely highly valued feature. Surprisingly, many respondents confessed their negative attitudes towards all kinds of chatbots. Our results also support the earlier findings that the major limitations of current implementations are lack of personalization [29], and users seem to trust less to chatbots in more complicated tasks [30].

Table 3. Strengths and weaknesses of the chatbot

Strengths	Frequency	Weaknesses	Frequency
Service always available	63%	I do not like chatbots	42%
No queues (can serve all students simultaneously)	38%	Does not encourage or inspirate me to study	37%
Quite easy to use	15%	Lack of personalization	29%
Has useful information to help course selection	12%	Do not know see other students' selections	21%

4.4 The evaluation of the development team

We can examine chatbot success and failure from two different angles. First, we can analyze it from the user perspective, and our results in the previous section exactly do that by indicating how users experienced Vivian. The second perspective is to analyze the chatbot provider's experiences.

Table 4. Development team's evaluations

Metric	Evaluation	Comments
We had enough resources	Strongly disagree	Not dedicate resources based on voluntarism
We had a clear business case	Strongly disagree	Not thought at all during the project
We had a clear use case	Somewhat agree	This is real problem for students
We had no legal or security problems	Strongly agree	Chatbot did not identify users or collected any user data
We were aware of user expectations	Neither agree nor disagree	Team did not have detailed knowledge of expectations
We provided high quality content	Somewhat disagree	Same content already available in other channels

Because success and failure are two sides of a coin, we decided to analyze the service provider's opinions with the six metrics that have been identified as the main reasons why chatbots fail in practice. These dimensions are the availability of resources, the strength of the business case, correctness of the use case, legal or security challenges, awareness of user expectations, and quality of the content [47]. We organized a workshop for the project team. During the workshop, team members first shared their views, and then they graded each dimension with a five-point agree/disagree scale with the following values: strongly disagree, somewhat disagree, either agree nor disagree, somewhat agree, and strongly agree. The results in Table 4 clearly point out that the main challenges of the project were related to the resources and the business case. The project did not have any external funding and was solely based on the shared interest of the project team. Also, the chatbot project was set up without detailed benefit and cost analysis or properly evaluating the possible operational changes. However, the project team saw that the use case was clear (students really need more help while planning their studies), and Vivian did not have any legal or security challenges because it did not identify the users nor collected any data from them.

5 Conclusions

In this paper, we aimed to analyze the suitability of the chatbots or virtual advisers in student counselling. Earlier studies have provided some mixed results, and therefore, this topic clearly required further studies. Our experiment revealed that modern IT tools offer a solid framework to create chatbots for different kinds of tutoring and counselling services. However, the results of our study revealed that developing a useful chatbot is a complex task with many dimensions.

The results of our study indicate that students found Vivian quite easy to use, but they were not as satisfied to it and the difference was statistically significant ($p = 0.05$). The negative NPS score (–38%) also clearly points out that Vivian is unlikely to be a viral hit among our students. That is a minor disappointment because the chatbot can remarkably reduce the need of human tutoring only if its widely accepted.

What could be the reasons behind our moderate success? Unfortunately, we cannot provide full answers but only provide some possible alternatives. Firstly, the qualitative part of the study indicated that chatbots could improve student counselling, and the main advantages of the chatbots were unlimited service hours and scalability. They can work 24/7 and can provide services to many students simultaneously. On the other hand, the main shortcomings of our chatbot were the minor significance of the individuality and the lack of inspiring effect. It was also interesting to recognize that many respondents had general negative attitude towards chatbots and automated services.

Secondly, there were without any doubt also some flaws in our implementation. When Vivian's chatbot specific characteristics were analyzed, the conversational capabilities got the lowest score, and this finding was statistically significant ($p = 0.05$). Thus, we must pay special attention to the communication capabilities of the chatbot in the future. Thirdly, students base their course selection to multiple factors like learning value, lecturer, course difficulty, prerequisite knowledge, and comfortability [31]. It is important to recognize that some of this information is not official and well documented, but it is only shared in informal contexts. Also, the findings of the earlier study [48] have indicated that "the occasional lack of adequate answers does not necessarily produce a bad experience, as long as the chatbot offers an easy path for follow-up with human customer service representatives." Both the transfer of unofficial silent information and the smooth transition from chatbot to human counsellors are important development areas in the future.

In addition to the users' point of view, we analyzed Vivian's pros and cons from the service provider's perspective. The development team pointed out that Vivian was doing a right thing (strong use case). On the other hand, the team learned that the development of a functional chatbot requires quite a lot of work (enough resources) and that its outcomes and changes to the counselling processes (business case) must be analyzed thoroughly before implementation.

All in all, the feedback clearly pointed out that Vivian is still a work in progress. To make it more useful, we must address the issues identified in this study. In the next version, Vivian must be connected better to other information systems of the university, and it should offer more flexible ways to share peer information and recommendations. Even after these modifications, we do not believe that Vivian will replace the need for human tutoring. Study planning is a complex decision-making process where students need the insight and advice from the experts, in addition to facts.

6 References

[1] P. B. Brandtzæg and A. Følstad, "Chatbots: Changing User Needs and Motivations", *Interactions*, vol. 25, no. 5, Sept.–Oct., pp. 38–43, 2018. https://doi.org/10.1145/3236669

[2] O. Zawacki-Richter, V. I. Marín and M. Bond, "Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where are the Educators?", *Int J Educational Technology Higher Education*, vol. 16, no. 39, 2019. https://doi.org/10.1186/s41239-019-0171-0

[3] J. K. Schulenberg and M. J. Lindhorst, "Advising is Advising: Toward Defining the Practice and Scholarship of Academic Advising", *NACADA Journal*, vol. 28, no. 1, 43–58, 2008. https://doi.org/10.12930/0271-9517-28.1.43

[4] S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Titnerger and H. Drachsler, "Are We There Yet? – A Systematic Literature Review on Chatbots in Education", 15 July, *Frontiers in Artificial Intelligence*, 2021. [online serial]. Available: https://www.frontiersin.org/articles/10.3389/frai.2021.654924/full

[5] S. Bhatnagar et al. "Mapping Intelligence: Requirements and Possibilities", in *Philosophy and Theory of Artificial Intelligence*, V. Müller, Eds. Cham. Springer, 2018. https://doi.org/10.1007/978-3-319-96448-5_13

[6] P. Wang, "On Defining Artificial Intelligence", *J. of Artificial General Intelligence*, vol. 10, no. 2, pp. 1–37, 2019. https://doi.org/10.2478/jagi-2019-0002

[7] C. Manning, "Artificial Intelligence Definitions", Human-Centered Artificial Intelligence, Stanford University, [online]. Available https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf [Access Sept. 1, 2022].

[8] J. E. Korteling et al., "Human-versus Artificial Intelligence", 25 March, 2021. *Frontiers in Artificial Intelligence*, 2021. [online serial]. Available: https://www.frontiersin.org/articles/10.3389/frai.2021.622364/full

[9] L. Gkinko and A. Elbanna, "The Appropriation of Conversational AI in the Workplace: A Taxonomy of AI", *Int. J. of Information Management*, August 2022. https://doi.org/10.1016/j.ijinfomgt.2022.102568

[10] J. Gao, M. Galley and L. Li, "Neural Approaches to Conversational AI", *Foundations and Trends® in Information Retrieval*, vol. 13, no. 2–3, pp. 127–298, 2019. https://doi.org/10.1561/1500000074

[11] M. Nuruzzaman and O. K. Hussain, "IntelliBot: A Dialogue-Based Chatbot for the Insurance Industry", *Knowledge-Based Systems*, vol. 196, May 21, 2020. https://doi.org/10.1016/j.knosys.2020.105810

[12] Drift, "2020 State of Conversational Marketing", Sept. 2020. [Online]. Available: https://www.drift.com/blog/state-of-conversational-marketing/ [Accessed Aug. 10, 2022].

[13] L. Zhou, J. Gao, D. Li and H-Y. Shum, "The Design and Implementation of XiaoIce, an Empathetic Social Chatbot", *Computational Linguistics*, vol. 46, no. 1, pp. 53–93, 2020. https://doi.org/10.1162/coli_a_00368

[14] J. J. Prochaska et al. "A Therapeutic Relational Agent for Reducing Problematic Substance Use (Woebot): Development and Usability Study", *J. of Medical Internet Research*, vol. 23, no. 3, 2021. https://doi.org/10.2196/24850
[15] S. Xu, "Microsoft Chatbot Spinoff Xiaoice Reaches $1 Billion Valuation", *Bloomberg.com* July 2021. [Online.] Available https://www.bloomberg.com/news/articles/2021-07-14/microsoft-chatbot-spinoff-xiaoice-reaches-1-billion-valuation [Accessed Sept. 1 2022].
[16] R. Dale, "The Return of the Chatbots", *Natural Language Engineering*, vol. 22, no. 5, pp. 811–817, 2016. https://doi.org/10.1017/S1351324916000243
[17] R. Meyer von Wolff, S. Hobert and M. Schumann, "How May I Help You? – State of the Art and Open Research Questions for Chatbots at the Digital Workplace." In Proc. the Hawaii International Conference on System Sciences '19, 2019. https://doi.org/10.24251/HICSS.2019.013
[18] D. Coniam. "The Linguistic Accuracy of Chatbots: Usability from an ESL Perspective", *Text & Talk*, vol. 34, no. 10, 2014. https://doi.org/10.1515/text-2014-0018
[19] J. Feine, S. Morana and A. Maedche. "Designing a Chatbot Social Cue Configuration System." In Proc. The International Conference on Information Systems '19, 2019.
[20] T. Rietz, I. Benke and A. Maedche. "The Impact of Anthropomorphic and Functional Chatbot Design Features in Enterprise Collaboration Systems on User Acceptance." In Proc. of the 14th International Conference on Wirtschaftsinformatik, 2019.
[21] G-J. Hwang and C-Y. Chang. "A Review of Opportunities and Challenges of Chatbots in Education." *Interactive Learning Environments*", July pp. 1–14, 2021. https://doi.org/10.1080/10494820.2021.1952615
[22] M. Sanji, H. Behzadi and G. Gomroki. "Chatbot: An Intelligent Tool for Libraries", *Library Hi Tech News*, vol. 39, no. 3, pp. 17–20, 2022. https://doi.org/10.1108/LHTN-01-2021-0002
[23] X. Sanchez-Dias et al. "A Knowledge-Based Methodology for Building A Conversational Chatbot as an Intelligent Tutor." In Proc. Mexican International Conference on Artificial Intelligence. Springer, Cham, 2018. p. 165–175. https://doi.org/10.1007/978-3-030-04497-8_14
[24] H-D. Truong et al. "Development of an AI Chatbot to Support Admissions and Career Guidance for Universities", *International Journal of Emerging Multidisciplinary Research*, vol. 4, no. 2, pp. 11–17, 2020. https://doi.org/10.22662/IJEMR.2020.4.2.011
[25] G.-J. Hwanga and C.-Y. Chang. "A Review of Opportunities and Challenges of Chatbots in Education", *Interactive Learning Environments*, July, pp. 2–14, 2021. https://doi.org/10.1080/10494820.2021.1952615
[26] D. Song, E. Y. Oh and M. Rice, "Interacting with a Conversational Agent System for Educational Purposes in Online Courses," In Proc. 10th International Conference on Human System Interactions (HSI), 2017, pp. 78–82. https://doi.org/10.1109/HSI.2017.8005002
[27] S. Kowalski, R. Hoffman, R. Jain and M. Mumtaz. "Using Conversational Agents to Help Teach Information Security Risk Analysis". In Proc. of the First International Conference on Social Eco-Informatics. 2011.
[28] P. Bii. "Chatbot Technology: A Possible Means of Unlocking Student Potential to Learn How to Learn", *Educational Research*, vol. 4, no. 2, pp. 218–22, 2013.
[29] P. Stone et al. "Artificial Intelligence and Life in 2030". One Hundred Year Study on Artificial Intelligence (Report of the 2015–2016 Study Panel), Stanford, CA: Stanford University, 2016.
[30] T. Nadarzynski, O. Miles, A. Cowie and D. Ridge. "Acceptability of Artificial Intelligence (AI)-led Chatbot Services in Healthcare: A Mixed-methods Study." *Digital Health*, Aug. 2019. https://doi.org/10.1177/2055207619871808
[31] E. Babad and A. Tayeb. "Experimental Analysis of Students' Course Selection", *British Journal of Educational Psychology*, vol. 73, pp. 373–393, 2003. https://doi.org/10.1348/000709903322275894

[32] J. Castells, P.-D. Mohammad, L. Galárraga, G. Méndez and M. Ortiz-Rojas. "A Student-Oriented Tool to Support Course Selection in Academic Counseling Sessions". In Proc of the 15th European Conference on Technology Enhanced Learning. 2020.
[33] A. Assiri, A. AL-Malaise and H. Brdesee. "From Traditional to Intelligent Academic Advising: A Systematic Literature Review of e-Academic Advising", *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 4, 2020. https://doi.org/10.14569/IJACSA.2020.0110467
[34] M. G. Méndezand and G. Arguello. "Best Practices of Virtual Advising: The Application of an Online Advising Portal", *FDLA Journal*, vol. 5, no. 1, 2020.
[35] M. Koivisto. *Mobile Information System Adoption and Use: Beliefs and Attitudes in Mobile Context*, Helsinki University of Technology, Finland, 2009.
[36] E. Rogers. *Diffusion of Innovations (4th ed.)*, New York, USA: The Free Press. 1995.
[37] F. Davis. "Perceived Usefulness, Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989. https://doi.org/10.2307/249008
[38] M. Fizbein and I. Ajzen. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, USA: Addison-Wesley. 1975.
[39] Y. J. Kim. "Social Influence Process in the Acceptance of a Virtual Community Service", *Information Systems Frontiers*, July. 2006.
[40] J. Nielsen, *Usability Engineering*, Boston, USA: Academic Press 1993. https://doi.org/10.1016/B978-0-08-052029-2.50009-7
[41] X. Faulkner. *Usability Engineering,* New York, USA: Palgrave. 2000. https://doi.org/10.1145/343048.343076
[42] B. Wixom and P. Todd. "A Theoretical Integration of User Satisfaction and Technology Acceptance", *Information Systems Research*, vol. 16, no. 1, pp. 85 – 102, 2005. https://doi.org/10.1287/isre.1050.0042
[43] W. Bleuel. "CSAT or CES: Does It Matter?" *Graziadio Business Review*, vol. 22, no. 1, 2019.
[44] M. Dixon, K. Freeman and N. Toman. "Stop Trying to Delight Your Customers", *Harvard Business Review*, Jul–Aug, 2010.
[45] F. Reichheld. "The One Number You Need to Grow", *Harvard Business Review*, Dec-, 2003.
[46] IBM, Watson Assistant: The Shortcut to Great Customer Service. White paper. IBM. 2016.
[47] A. Janssen, L. Grützner and M. Breitner. "Why do Chatbots Fail? A Critical Success Factors Analysis". In Proc. International Conference on Information Systems (ICIS) 2021.
[48] A. Følstad and M. Skjuve. "Chatbots for Customer Service: User Experience and Motivation." In Proc. The 1st International Conference on Conversational User Interfaces. 2019. https://doi.org/10.1145/3342775.3342784

7 Author

Matti Koivisto is a Principal Lecturer at South-Eastern Finland University of Applied Sciences.

This is a revised and extended version of a presentation given at The Learning Ideas Conference 2022, held in New York and online June 15–17, 2022. Article submitted 2022-09-20. Resubmitted 2023-02-09. Final acceptance 2023-02-09. Final version published as submitted by the authors.