Examining the Design Space of AI-Mediated Social Matching Among Online Learners

Qiaosi Wang, Ida Camacho, Ashok K. Goel

To appear in Social and Emotional Learning - An Inclusive Learning Analytics Perspective,
edited by Elle Wang et al, Springer, 2021

Abstract Online education has been growing in demand over the years. However, online learners frequently experience social isolation, which negatively impacts their learning experience and outcome. In this chapter, we investigate the design space of social matching systems to help foster social connections among online learners. Specifically, we seek to answer three core design questions: (1) What data should be collected? (2) How to design technology to support students' interactions with one another? (3) What are students' concerns about the ethics of AI-mediated social matching? We begin by exploring the feasibility, design, and concerns of AI-mediated social interactions through existing literature. We then present our ongoing work on the design and use of AI conversational agents as social matching systems in the online learning context. Finally, we outline future directions for research on designing human-centered social matching systems in online learning.

1 Introduction

With growing demand for online for-degree programs as well as non-degree courses, online learning has become critical in shaping the landscape of education. The success of online learning however depends on multiple factors, one of which is the degree of social connectedness among online learners (Aldosemani et al., 2016; Arbaugh et al., 2008). Strong social ties among online learners are crucial to raise their degree of satisfaction (Hostetter and Busch, 2006; Rovai, 2001), reduce dropout rates (Rovai, 2002), and stimulate intellectual exchange by providing a safe atmo-

Qiaosi Wang
Georgia Institute of Technology e-mail: qswang@gatech.du

Ida Camacho
Georgia Institute of Technology e-mail: icamacho@gatech.edu

Ashok K. Goel
Georgia Institute of Technology e-mail: ashok.goel@cc.gatech.edu

sphere (Aldosemani et al., 2016; Rovai, 2001). Many learning scientists consider learners' social presence an integral part of their success in online learning (Garrison and Arbaugh, 2007; Arbaugh et al., 2008; Lave et al., 1991).

However, the social dimension of online learning has not received as much attention as the cognitive aspect. Much research in online learning has been devoted to investigating methods and techniques that can improve online learners' cognitive processes and behaviors. For example, questions on how to improve student engagement (Wang et al., 2020c; Ou et al., 2019), learning achievement (Rohloff et al., 2020), teaching effectiveness (Patikorn and Heffernan, 2020) have been the major focus for many online learning researchers. Yet strong social bonds between students are often the basis of optimal learning processes and experiences. For learners to be open to making mistakes and for them to willingly exchange ideas, they need to have a certain level of trust in each other, feel a sense of social belonging in the learning community, and feel close to each other for more risk-taking and adventurous learning attitudes (Kreijns et al., 2003; Aldosemani et al., 2016). Yet little research has examined design of information technology to support online learners' social interaction processes.

One promising way to facilitate online learners' social interaction process is to help online learners develop affinity for one another through the discovery of shared identity (Sun et al., 2019; Wang et al., 2020b). This discovery of shared identity can be facilitated through the use of **social matching systems**. A social matching system is a particular type of recommender system that aims at providing recommendations of people that might be of interest for someone to connect with (Terveen and McDonald, 2005; Mayer et al., 2015). Social matching systems, while prominently used in the context of online dating (Zytko et al., 2018), have also been employed to rediscover old friends on social networks (Motoyama and Varghese, 2009; Chen et al., 2009), link job seekers with potential employees (Olsson et al., 2020) and connect academic researchers to local community collaborators (Zytko and DeVreugd, 2019).

Many social matching systems follow a five-stage process: profiling users, computing matches, introduction, interaction, and feedback (Terveen and McDonald, 2005). To help support online learners' social interaction process, a social matching system must first build a *profile* of each learner through collecting relevant data and information that could be useful in finding matches, potentially based on the learner's background, geographical location, interests and hobbies, classes taken, progress in the course and the program, etc. Using this profile, the system could *compute matches* for the learner based on some criteria either explicitly set by the student (e.g., want to connect with students located in the same city) or implicitly inferred by the system (e.g., connect students who are going through the same learning modules). After the matches are computed, the system should *introduce the matches* together in some form, for instance, directly putting matches into contact or providing the matches' contact information to one another. Depending on the learners and the learning context, the system could also intervene during learners' *interaction* processes, for example, post ice breaker questions to help them start a conversation. Finally, given that learners' profile might change over time or they might not like the recommended matches, they should be able to provide *feedback* to the system to optimize future matches.

While this basic process offers a general model of how a social matching system could operate in an online learning context, like the design of any other information technology, more detailed design requirements are needed to create actual technology that can tailor to learners' preferences and needs while also addressing privacy and other ethical considerations. *Thus the core question that we seek to explore in this book chapter is how to design social matching systems for enhancing social interactions among online learners from a human-centered perspective.* To investigate this question in the context of the five-stage process described above, we break it down into three sub-questions and map the sub-questions to different stages of the process as illustrated in Fig. 1:

- What kinds of data should be collected to help online students make social connections?
- How to design social matching systems to support the processes of social interaction among online learners?
- What ethical concerns might students have for social matching systems in online learning environments?

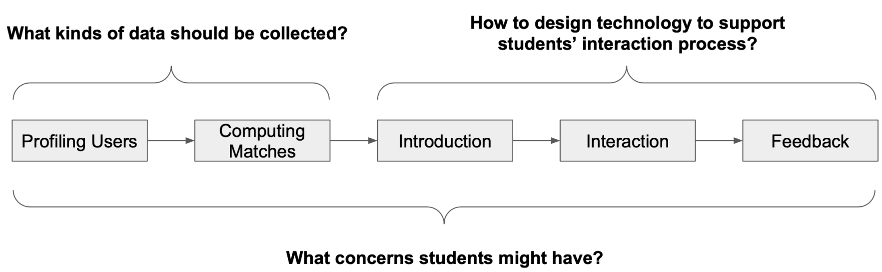

Fig. 1 Three design questions to be explored on data, interaction, and concerns at different stages of social matching process in online learning context.

To explore these questions, we first draw from relevant work in a variety of fields related to learning analytics and Computer-Supported Cooperative Work (CSCW) to identify relevant design implications of social matching systems in online learning environments. Building upon existing literature, we present our ongoing work and latest findings on the design and development of a social matching system in the context of Georgia Tech's Online Master of Science in Computer Science (OMSCS) program (Galil, 2020; Joyner et al., 2019). While social matching systems can take many forms, we specifically focus on using Conversational Agents (CAs) as social matching systems given CAs' success in providing emotional and social support within online communities (Nordberg et al., 2019; Narain et al., 2020). This work builds on but is different and separate from our earlier work on the AI teaching assistant named Jill Watson (Goel and Polepeddi, 2016, 2019): while Jill answered learners' questions on discussion forums of online classes and thereby enhanced teacher presence, the present work addresses the issue of promoting social interactions among the online learners. Taking a human-centered design perspective, we outline the design space of AI-mediated social matching systems in online learning environments based on empirical evaluation and deployment of CAs as social matching systems. We then highlight future directions of designing AI-mediated social matching systems in online learning environments.

2 Related Work

In this section, we review relevant work to understand the design space for AI-mediated social matching systems in online learning environments. Following our three questions mapped out to the basic model of social matching systems (Figure 1), we first discuss several crucial elements of establishing interpersonal connections and how these elements could be highlighted and inferred from students' digital footprints. Next, drawing upon theoretical frameworks from CSCW, we highlight the design characteristics of AI technology that can support remote social interaction processes. Finally, drawing from the well-known ethical challenges and concerns in the use of AI technologies powered by users' online data, we discuss the potential concerns social matching systems might raise for online learners.

2.1 Profiling and Computing Matches for Online Learners

While establishing connections online can be very different from and more challenging than in in-person context, many elements of interpersonal connections are shared across both settings. Social psychologists have been studying interpersonal attraction for decades and identified different types of factors that are crucial to building an amicable interpersonal relationship (Terveen and McDonald, 2005): personal characteristics (e.g., personal preferences, personality), demographics (e.g., gender, profession), and familiarity (e.g., time spent together). Based on prior research, people are more prone to connect with those who share similarities in personal characteristics and demographics (Granovetter, 1973; Gilbert and Karahalios, 2009), as well as high levels of familiarity (Kraut et al., 1988). On top of establishing interpersonal connections, cooperative actions also require that the individuals are likely to meet again in the future, the individuals can identify each other prior to interactions, and the individuals possess adequate amount of information of people's past behaviors (Terveen and McDonald, 2005; Kollock, 1997).

Perhaps one of the biggest challenges for online social interactions is that most people find that inferring this kind of insight is difficult, sometimes impossible, based on solely on another person's online behavior (Kehrwald, 2008). While in in-person interactions, most people can efficiently, and often accurately, gauge people's personality, characteristics, and age based on their appearances and behaviors, online environment is often text-based, stripping people's ability to make inferences about another person's demographics and other characteristics.

Yet online environments also present many opportunities for AI agents to make such inferences: it is easier to capture and retain information in online environments than in in-person contexts, which can enable AI agents to make inferences from people's online behaviors. For example, based on people's digital footprints in online environments, researchers were able to use AI techniques to infer people's mental states (e.g., stress) from online forum data (Saha and De Choudhury, 2017; De Choudhury et al., 2013), predicting people's personality from social media cues (Skowron et al., 2016; Farnadi et al., 2016), and inferring about people's interpersonal ties using social media data (Gilbert and Karahalios, 2009). These studies all point to the feasibility of compensating for the lost social cues in online social interactions using people's digital footprint.

Leveraging people's online information to infer behaviors and personal characteristics is hardly an uncharted area in online education— the field of learning analytics and educational data mining have been analyzing online learners' data to make inferences about students for many years (Du et al., 2021; Avella et al., 2016). The overall objectives for learning analytics is to leverage online learners' data to predict learner performance, offer decision support for teachers and learners, detect behavioral patterns and learner models, as well as predict dropout rates (Du et al., 2021; Avella et al., 2016). To accomplish these objectives, researchers have been able to leverage many data sources readily available in online learning —students' educational records, demographics, textual data of online discussions, facial expression, frequency of logins, duration of content accessed— that can shed light on students' learning progress, learning patterns, learning behaviors, etc (Du et al., 2021). However, these efforts at using learning analytics for enhancing online learning have mostly focused on the cognitive aspects of learning; the potential of using learning analytics approaches to support students' social interaction process requires further exploration.

2.2 Designing Technology-Mediated Remote Social Interactions

Decades of CSCW research has produced several well-established theoretical frameworks to guide the design of technologies in supporting remote interactions. Among these theoretical frameworks, both Ackerman's social-technical gap (Ackerman, 2000) and Erickson and Kellogg's social translucence (Erickson and Kellogg, 2000) draw inspiration from in-person social interactions to design technology that can support remote social interactions.

Ackerman defines social-technical gap as *"the great divide between what we know we must support socially and what we can support technically"* (Ackerman, 2000). In his seminal work, Ackerman points out that when technology mediates remote interactions, they are often designed to be rigid, reductionist, and do not allow sufficient ambiguity compared to in-person interactions (Ackerman, 2000). Much research has since adopted this framework and identified the social-technical gap in a variety of contexts such as health tracking (Chung et al., 2017), collaboration among telesurgery teams (Duysburgh et al., 2014), and online collaborative consumption (Gheitasy et al., 2015). In his original piece, Ackerman proposed first-order approximations — solutions that partially solve the problem but with known trade-offs — to help bridge the social-technical gap. One optimal approximation is to design augmentative information technology, for example, by offering advice to users (Ackerman, 2000). This potentially can be accomplished through the use of CAs (Lee et al., 2017).

While the idea of social-technical gap typically acts as a general guide and call-to-action in CSCW research, to bridge this gap between social and technical requirement, Erickson and Kellogg go a step further and outline detailed principles on designing towards socially translucent systems to support natural online interactions (Erickson and Kellogg, 2000). Specifically, socially translucent systems have three characteristics: visibility, awareness, and accountability. *Visibility* refers to system's ability of making social information more visible; *Awareness* refers to people's ability to know each others' existence; *Accountability* refers to system's ability to hold people accountable for their behavior by generating and enforcing social rules. Erickson and Kellogg posit that these three characteristics allow people to observe, imitate, aware, and interact with others socially in in-person context, and thus building socially translucent system is a fundamental requirement for people to carry out normal interactions online (Erickson and Kellogg, 2000). Since then, social translucence has been often employed in the design of technology-mediated interactions. For example, prior research has developed methods to support collective awareness through creating common repository to generate mutual understanding for members of globally distributed teams (Bjørn and Ngwenyama, 2009) and conducting synchronous coding sessions for learner engagement (Byun et al., 2020).

In summary, this body of work emphasizes the lack of naturalness in remote interactions compared to in-person interactions — social-technical gap — and how technology can be designed to be socially translucent in order to create the naturalness in online environment. While these two theoretical frameworks have not been widely used in research on online learning, several studies on online learners' social presence and social interactions have offered some support to the generalizability of these issues in the online learning context: the lack of visibility of social cues (Kehrwald, 2008; Sun et al., 2019) and the diminished accountability and motivation in reaching out to others (Kehrwald, 2008) all contribute to students' feeling of social isolation in online learning environments.

2.3 Potential Challenges in Social Matching Among Online Learners

Like many other AI systems that leverage big data, social matching systems and learning analytics approaches present several potential ethical challenges and concerns. Well known ethical concerns such as privacy, consent, anonymity, and accuracy of data are shared across the use of social matching systems and interventions based on learning analytics (Terveen and McDonald, 2005; Avella et al., 2016; Wang et al., 2020c).

Both social matching systems and learning analytics interventions produce results based on either user data that are voluntarily offered by the users or user data that are available but not explicitly consented to by the users such as postings on a public forum. Given that humans are social creatures, one ethical dilemma social matching systems face is the fact that many people often are okay with their sensitive personal information being used in specific contexts — and sometimes even voluntarily offer it — but this can lead to oversharing (Terveen and McDonald, 2005). For example, sensitive information such as student grades is commonly collected in learning analytics approaches, often in order to assess students' learning progress. However, it remains questionable whether students are aware of the extent to which their data is being collected and analyzed in online learning environments since usually only instructors and institutions have access to the data and the results (Slade and Prinsloo, 2013). Designing social matching systems in online learning environment thus would require transparency of the processes of data collection and analysis, as well as careful informed consent procedures to address privacy and ethical concerns.

One common pitfall for AI systems that are powered by big data is the fact that sometimes individuals are treated more like data points than humans with identity and agency. One important characteristic to keep in mind of is that people's identity is often transient and temporal (Slade and Prinsloo, 2013; Terveen and McDonald, 2005): students' learning behaviors as well as their preferences can change over time. Feedback from the students regarding their social matches thus can play a crucial role for the system to update and caliberate future recommendations (Terveen and McDonald, 2005). In learning analytics approches, treating students as agents could mean asking for their collaboration throughout the analytical process (Slade and Prinsloo, 2013). This not only means performing data collection, analysis, and usage only with students' explicit and specific consent, but also to ensure that the system can leverage information students voluntarily offer to help them achieve their own learning goals (Slade and Prinsloo, 2013).

The most basic functionality of social matching systems is to recommend and match people with similarities (Terveen and McDonald, 2005), which is based on people's natural similarity-seeking behaviors during in-person interactions (Olsson et al., 2020). In online environment this tendency towards similarity still persists—prior research has found that in online team formation, people tend to team up with those who are similar to them and thus lead to non-diverse and segregated teams (Gómez-Zará et al., 2019). Previous research has pointed out that this fundamental design characteristic of social matching systems can lead to ethically concerning consequences such as the creation of echo chambers and polarization in the community (Olsson et al., 2020). Olsson et al. further argue that recommendation systems should not encourage biased human behaviors and that one potential solution is to enable social serendipity and random encounters in online social matching (Olsson et al., 2020).

2.4 Summary

To understand the design space of social matching systems in online learning contexts, we first reviewed relevant literature to explore the three core design questions about data, interactions, and concerns (see Figure 1). Based on the existing literature, we found that understanding and identifying similarities in personal characteristics, demographic, and familiarity is crucial in establishing social connections, both in-person and online (Terveen and McDonald, 2005; Granovetter, 1973). In online contexts, recent development in AI and Natural Language Processing (Saha and De Choudhury, 2017; Du et al., 2021; Avella et al., 2016) allow fairly accurate inference of such social information and thus should be leveraged to collect relevant data in profiling and computing matches when designing social matching systems in online learning. To design social matching systems that can support students' social interactions, principles of social translucence (Erickson and Kellogg, 2000) and social-technical gap (Ackerman, 2000) could be applied to help replicate the naturalness of in-person interactions to online environment. Concerns regarding AI-mediated social matching systems could arise at any stage of the social matching process, specifically, privacy, consent (Slade and Prinsloo, 2013), oversharing of personal information (Terveen and McDonald, 2005), updating students' transient identities (Slade and Prinsloo, 2013; Terveen and McDonald, 2005), and the unintended creation of echo chambers (Olsson et al., 2020) are all valid concerns and should be taken into account when designing social matching systems for online learners.

Based on these design considerations that we identified through existing literature, we designed and deployed a community-facing CA called SAMI (Social Agent Mediated Interaction) (Goel, 2020) to perform social matching among online students. In the next section, we describe our design and deployment of SAMI in an online learning context.

3 SAMI: Conversational Agents as Social Matching Systems

Due to its human-like characteristics and the ability to converse with people, CAs have been widely used to provide social and emotional support in both dyadic interactions and community contexts. Prior research has demonstrated the positive effect of using CAs to facilitate mental health patients' self-disclosure during consultations (Lee et al., 2020), help healthcare professionals manage occupational stress (Yorita et al., 2020), and provide social support to older adults who are socially isolated (Simpson et al., 2020; Ring et al., 2015). Designing social matching systems as CAs is thus a promising way to support online learners who feel socially isolated yet also requires more design explorations.

Inspired by the prior research's usage of textual data in online discussion forum, SAMI leverages students' self-introduction posts on the discussion forum, where online learners usually conduct class-related discussions and posting self-introductions at the beginning of the semester. Specifically, SAMI utilizes Natural Language Processing to extract different entities such as hobby, city, country from students' self-introduction posts in order to build a profile for each online student. Online students can opt-in to receive SAMI response by adding "#ConnectMe" in their introduction post as seen in Fig. 2.

The current version of SAMI matches students shared similarities among students such as proximate geographical locations and similar hobbies, etc. (see Fig. 2) After identifying students' preferred matching criteria, SAMI creates a private group of all students with commonalities and then invites each student to the private group. To further engage students in building connections, SAMI also posts ice-breaker questions within each group (see Fig. 3).

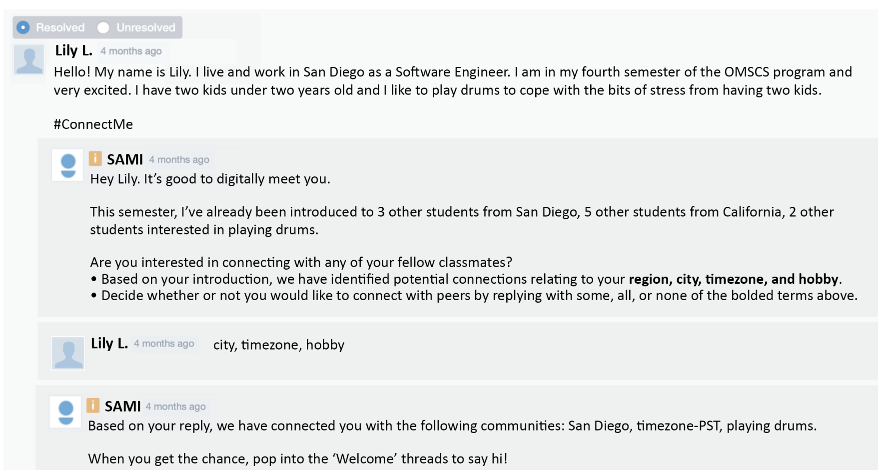

Fig. 2 SAMI uses NLP to extract entities from student's post, inquires about student's matching criteria, then puts the student in private groups with other students who share similarities.

We constructed the first basic version of SAMI in 2017 for Georgia Tech OMSCS class on Knowledge-Based AI (Goel and Joyner, 2017, 2016). Since then we have both incrementally enhanced SAMI's capabilities and deployed it in additional OMSCS classes (Goel, 2020). We have also conducted detailed evaluations and collected extensive student feedback on SAMI for future improvements in its design and delivery. W present our evaluation studies and findings in the next section.

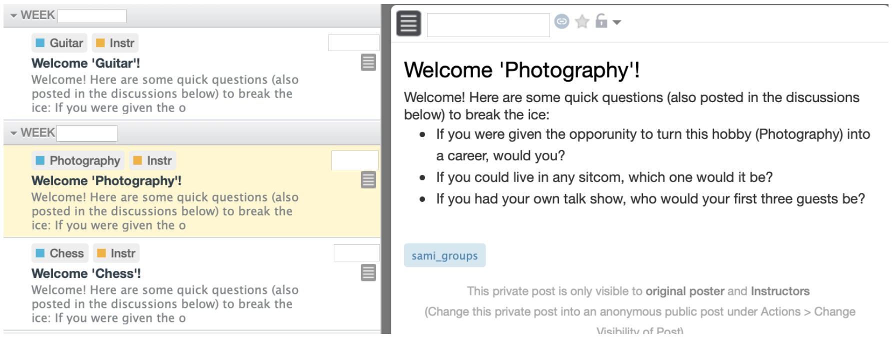

Fig. 3 SAMI creates different private groups based on entities identified from students' introduction posts and then put students with similarities in the according groups (e.g., similar hobby). SAMI also posts ice breaker questions in each group to help students start the conversation.

4 Evaluation of A Social Matching System in Online Learning

In this section, we present and discuss our findings from a survey study (Wang et al., 2020b) and a set of semi-structured interviews (Wang et al., 2020a) with students in the OMSCS program regarding their existing challenges in building social connections with other students as well as their experience with SAMI. Based on our analysis, we outline design implications for building social matching systems in online learning environments, highlight online learners' concerns regarding the use of social matching systems in online learning, and discuss potential directions for designing social matching systems for online learners.

4.1 Designing Beyond Social Translucence in Social Matching Process

Designing AI systems that can support students' interaction with one another is a design problem that could manifest through the introduction, interaction, and feedback stages of the social matching process (see Fig. 1). This raises several questions: How much information about a given student should the system share with other students? How should the system introduce the matches to each other? How much should the system intervene during the introduction and interaction stages? In this section, we seek to answer these questions by understanding online learners' current challenges in remote social interactions and interpreting their feedback on SAMI in helping online learners connect with one another. We point out that the design of social matching systems should not only provide social translucence in the current online learning platforms, but it should also seek to bring the randomness and authenticity commonly seen in in-person interactions into online social interactions.

Making social signals visible among online learners. Reaching out and building connections with strangers can be an intimidating process. During in-person interactions, people are able to gain social cues from another person's behaviors or facial expressions. However, most of these social cues become invisible in an online environment (Erickson and Kellogg, 2000). In our study, we found that many online learners are hesitant in reaching out to other people due to the loss of social cues— they do not know whether other students are willing to connect with them or how their messages will be received by other students. After SAMI was deployed in participants' online classes, students were more readily able to identify social signals identified by SAMI. Some students interpreted the "#ConnectMe" in other students' self-introduction posts as a signal of students' willingness to build social connections. When designing social matching systems in online learning environment, making social signals visible is an important starting point and could be implemented easily by adding simple features. For example, one potential feature is adding icons on students' avatars to indicate students' willingness to connect with others to improve visibility of social cues (Szostek et al., 2008).

Raising awareness of potential social companions. One of the main goals and advantages of online learning is to help education scale by giving more students the opportunities to learn (Larreamendy-Joerns and Leinhardt, 2006). One result of this scaling is that online classes usually have hundreds or even thousands of students per class. The downside of having these many students per class is that the class size reduces students' awareness of other students' existence in the program/class, which negatively contributes to online students' social interaction process. This diminished awareness poses several challenges in the learners' social interaction processes. First, the overwhelming number of students and activities within each class made it extremely difficult for online students to identify individual students that they could potentially build social connections with. Second, with hundreds of students communicating via the class discussion forum and chat group, these main communication channels quickly became a wall of text, which makes most interactions there seemed impersonal. With the deployment of SAMI, participants became more conscious of other students who shared similar backgrounds, interests or experiences with them. Even without the personalized recommendations, students said that just scrolling through SAMI's replies to students self-introduction posts made them realize there were other students with different or similar experiences, which made the online learning environment seem more personable and personalized. Future social matching systems thus could raise online student's awareness of other students by highlighting students' shared identities. Offering statistics of the entire class or the program to demonstrate the diverse student population could also help students to increase the personable feelings in online learning environment.

Providing accountability to the social interaction process. Erickson and Kellogg pointed out that while awareness and acccountability often co-occur in physical world, they are not usually coupled in the online spaces (Erickson and Kellogg, 2000). Accountability is often fostered through the creation of social norms in a community that hold people accountable for their social behaviors (Erickson and Kellogg, 2000). In online learning environments, both the existence and the lack of social rules could prohibit online learners' social interactions. While the intention of the online class discussion forum is to replicate the physical classroom where students could have interactions and discussions about and beyond academics, the implicit social rule to use online class discussion forum only for academic discussions makes students feel accountable to only have academic discussions on the forum instead of casual conversations. Although online learners often get the chance to get to know fellow students through group projects, after the semester ends, they don't usually "encounter" one another anymore, which reduces students' feeling of accountability to talk with each other again. One feature of SAMI is to put students with similarities directly into a private group and post ice breaker questions to help students start the conversation. Participants in our study believed that by putting students directly in touch with each other, SAMI not only alleviated students' mental barrier in initiating the conversation, but also made students feel accountable to start building the connections because SAMI already "started" the conversations between students. Providing accountability in the social matching process thus could also take the form of the AI agent initiating the conversation between the matches or offering timely nudges for individuals to start building the connections.

Creating randomness and spontaneity in remote social interaction. While many challenges online students encounter during their social interaction process originate from the lack of social translucence in online learning environment—visibility, awareness, accountability— another major challenge we identified goes beyond social translucence and highlights the social-technical gap (Ackerman, 2000) in online social interactions. According to our participants, the randomness and spontaneity that were crucial and inherent in in-person environment currently could not be supported in the online learning environment. For example, in in-person educational environment, students can often randomly run into each other on campus or having work conversations that organically lead to more social activities. However, in online learning programs, the social and learning aspects typically are separated, especially when compared to traditional in-person educational programs. Instead of forming social connections organically during the process of taking classes or walking around campus that are inherently built into the on-campus educational experience, online learners have to establish social connections in a more intentional way (e.g., driving for an hour to attend a local meet-up with other online students). While highlighting shared identities or explicitly expressing social signals could help create social translucence into the online social interaction process, the nuances and subtlety of in-person interactions should also be preserved when performing remote social matching (Olsson et al., 2020; Ackerman, 2000). A potential direction for future designs of online social matching systems among online learners is to intentionally create seemingly random matches (e.g., students who seem very different at first but have "deep" similarities) or introduce matches in seemingly random encounters (e.g., introduce students who are reviewing the same lecture materials).

Summary of design implications. Designing social matching systems for online learners thus should focus on bringing social translucence (Erickson and Kellogg, 2000) into the existing online learning platforms by increasing visibility of social signals, raising awareness of potential social companions, and providing accountability to the social interaction process. However, our work also reveals the social-technical gap (Ackerman, 2000) in remote social interactions and suggests that current online learning environment lacks the randomness and spontaneity of in-person interactions. Echoing with prior research (Kreijns et al., 2003), our work provides further empirical evidence that *social interactions in online learning environment cannot, and should not, be taken for granted to naturally happen just because the platforms allow it*. Future research can explore the design question of how to replicate the randomness and spontaneity of in-person social interactions to online learning environments.

4.2 Towards Collaborative Social Matching In Online Learning

Throughout the model of the social matching system process (see Figure 1), the design questions regarding both data and interactions are targeted at specific stages: questions about data could manifest from profiling users and computing matches; and questions about designing interactions could exhibit in the introduction, interaction, and feedback stages. Concerns that online students have regarding social matching systems in online learning environment come from issues surrounding both data and interaction design. Based on students' feedback on SAMI, we found that students are concerned about losing agency during the social matching process when it is mediated by an AI agent. We also found that students prefer the agent to be more transparent about the matching process and mechanism. A little to our surprise, the students in our survey did not express many concerns regarding data privacy. Based on our findings, we propose the future direction of designing social matching process as a collaborative process between the system and the students to mitigate students' concerns regarding both data and interactions as discussed below.

Preference in system transparency. When asked about how SAMI could improve in the future, many participants wanted SAMI to be more transparent about its decision-making process and working mechanism for a smoother interaction between students and SAMI. This preference stems from students' belief that if they can better communicate with SAMI using similar vocabularies, the matching results could be more accurate. Even during the evaluation phase of SAMI, some students thought out loud and wondered why SAMI matched some of their hobbies successfully but not the other ones. In human-AI collaborative decision-making processes, transparency and the willingness to collaborate are crucial for a desirable collaborative experience and outcome (Cai et al., 2019). Fortunately, aligned with prior literature (Liao et al., 2020; Jhaver et al., 2019), participants in our study indicated their willingness to understand the AI agent's vocabulary beforehand to adjust their choice of words during communication in order to improve the accuracy of matching results. Echoing with prior research that urges designing collaborative learning analytics interventions (Slade and Prinsloo, 2013), designing collaborative social matching systems with online learners could also be a promising direction to offer more system transparency as well as ensuring the accuracy of matching results.

Concerns about losing agency in building social connections. One concern raised by some students about the prospect of continuing usage of SAMI in the online learning program is the possibility of losing agency in making decisions on building social connections with other online students. This concern was mostly based on SAMI's feature of directly putting students in private groups. This feature, while created an adequate amount of social pressure for students to start the initial conversation, was also critiqued by students who said they wanted more freedom in choosing which group they could join or whom they should connect with. Building social connections with others is inherently a very personal decision-making process. While social matching systems can provide convenience and efficiency in facilitating online learners' social interaction process by suggesting matches that students otherwise would not be able to find, we found that students are unwilling to cede control of the decision-making process in choosing whom they should connect with (Sundar, 2020). This would require the social matching system to work with the students collaboratively throughout the decision-making process to maintain students' sense of agency. This could be accomplished by having the system communicate with the students about all the progress that has been made in computing matches, asking students to set and revise the matching criteria, and incorporating students feedback in future matching computations. However, it is important to keep in mind that the social matching system should also provide an adequate level of social pressure and accountability, such as putting students directly into groups, for students to initiate the conversation. This design issue of balancing between maintaining student's agency in decision-making and creating social pressure and accountability in social matching process thus requires further exploration.

Concerns about privacy. Even though privacy is often a concern for AI systems that leverage public data (Fiesler et al., 2016), most participants in our survey did not express privacy concerns regarding SAMI. In fact, many online students indicated that they were willing to offer more information for SAMI to find more connections for them, which also aligns with prior work that users are often willing to provide information to social matching systems to get connected (Terveen and McDonald, 2005).In our study, SAMI only obtains public information presented on the online forum, which might have alleviated online students' privacy concerns—some students in the study believed that the goal of posting on public forum was for others to see it. To achieve high accuracy in social matching systems, accessing latent behavioral data that users don't explicitly consent to would be inevitable and might result in violations of user privacy (Fiesler and Proferes, 2018; Fiesler et al., 2016). This is especially important for social matching systems as users are often more likely to disclose sensitive information for more accurate matching results (Terveen and McDonald, 2005). When CAs perform social matching for students, privacy issue would require more scrutiny as CAs possess human-like characteristics that could encourage people's self-disclosure during conversations (Lee et al., 2020; Følstad et al., 2018; Ischen et al., 2019) which might lead students to unintentionally disclose sensitive information that could be used by the CA to improve matching accuracy. The balance between privacy and accuracy in CAs as social matching systems in online communities thus requires further exploration.

Directions for future designs. Based on students' preferences and concerns regarding the use of social matching systems in online learning, we highlight the design direction of designing towards a collaborative social matching process between the students and the social matching system. Collaborative social matching could not only offer transparency about the social matching process, potentially increase matching accuracy, but also could help mitigate students' concerns about losing agency in building social connections. While students in our study did not express privacy concerns, designers should be cautious about the use of both public and private information, especially users' tendency to overshare sensitive personal information to ensure matching accuracy. Our work also raise several design issues that future research should explore such as user vs. system control over decision-making in human-AI collaboration and balancing between maintaining user agency and creating social pressure in the social matching process.

5 Conclusions

As online learning is adopted by increasingly large number of educational institutions, the social dimension of online learning requires more attention from researchers in a variety of fields. Social matching systems as an information technology that can support online learners' social interaction process is a promising first step to help reduce online learners' feelings of social isolation. However, as with many data-driven AI approaches, the design space of social matching systems in online learning requires further examination to cater to online learners' challenges and needs in remote social interactions as well as in mitigating potential privacy and ethical concerns. In this chapter, we explored the design space of social matching systems through three core design questions about data, interactions, and concerns. Drawing upon relevant literature, we established the feasibility of inferring social information from online learners' digital footprints, discussed related theoretical frameworks in designing technology to support remote social interactions, and presented existing concerns and challenges in AI-mediated approaches that leverage student data. We further elaborated on this initial design space drawn from existing literature through a discussion of our ongoing work on the design and evaluation of an AI conversational agent as a social matching system in an online learning context. Based on our findings, we outlined the design implications of designing social matching systems to provide social translucence as well as to create randomness and spontaneity in remote social interactions. We then pointed out directions for future work on building collaborative social matching systems to mitigate students' concerns and potential challenges.

References

Ackerman MS (2000) The intellectual challenge of cscw: the gap between social requirements and technical feasibility. Human–Computer Interaction 15(2-3):179–203

Aldosemani TI, Shepherd CE, Gashim I, Dousay T (2016) Developing third places to foster sense of community in online instruction. British Journal of Educational Technology 47(6):1020–1031

Arbaugh JB, Cleveland-Innes M, Diaz SR, Garrison DR, Ice P, Richardson JC, Swan KP (2008) Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework using a multi-institutional sample. The internet and higher education 11(3-4):133–136

Avella JT, Kebritchi M, Nunn SG, Kanai T (2016) Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. Online Learning 20(2):13–29

Bjørn P, Ngwenyama O (2009) Virtual team collaboration: building shared meaning, resolving breakdowns and creating translucence. Information systems journal 19(3):227–253

Byun J, Park J, Oh A (2020) Cocode: Co-learner screen sharing for social translucence in online programming courses. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp 1–4

Cai CJ, Winter S, Steiner D, Wilcox L, Terry M (2019) " hello ai": Uncovering the onboarding needs of medical practitioners for human-ai collaborative decision-making. Proceedings of the ACM on Human-computer Interaction 3(CSCW):1–24

Chen J, Geyer W, Dugan C, Muller M, Guy I (2009) Make new friends, but keep the old: recommending people on social networking sites. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 201–210

Chung CF, Gorm N, Shklovski IA, Munson S (2017) Finding the right fit: understanding health tracking in workplace wellness programs. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp 4875–4886

De Choudhury M, Gamon M, Counts S, Horvitz E (2013) Predicting depression via social media. In: Proceedings of the International AAAI Conference on Web and Social Media, vol 7

Du X, Yang J, Shelton BE, Hung JL, Zhang M (2021) A systematic meta-review and analysis of learning analytics research. Behaviour & Information Technology 40(1):49–62

Duysburgh P, Elprama SA, Jacobs A (2014) Exploring the social-technological gap in telesurgery: collaboration within distributed or teams. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, pp 1537–1548

Erickson T, Kellogg WA (2000) Social translucence: an approach to designing systems that support social processes. ACM transactions on computer-human interaction (TOCHI) 7(1):59–83

Farnadi G, Sitaraman G, Sushmita S, Celli F, Kosinski M, Stillwell D, Davalos S, Moens MF, De Cock M (2016) Computational personality recognition in social media. User modeling and user-adapted interaction 26(2):109–142

Fiesler C, Proferes N (2018) "participant" perceptions of twitter research ethics. Social Media+ Society 4(1):2056305118763366

Fiesler C, Lampe C, Bruckman AS (2016) Reality and perception of copyright terms of service for online content creation. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, pp 1450–1461

Følstad A, Nordheim CB, Bjørkli CA (2018) What makes users trust a chatbot for customer service? an exploratory interview study. In: International Conference on Internet Science, Springer, pp 194–208

Galil Z (2020) Omscs: the revolution will be digitized. Communications of the ACM 63(8):27–29

Garrison DR, Arbaugh JB (2007) Researching the community of inquiry framework: Review, issues, and future directions. The Internet and higher education 10(3):157–172

Gheitasy A, Abdelnour-Nocera J, Nardi B (2015) Socio-technical gaps in online collaborative consumption (occ) an example of the etsy community. In: Proceedings of the 33rd Annual International Conference on the Design of Communication, pp 1–9

Gilbert E, Karahalios K (2009) Predicting tie strength with social media. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 211–220

Goel A (2020) Ai-powered learning: Making education accessible, affordable, and achievable. arXiv preprint arXiv:200601908

Goel A, Joyner D (2016) An experiment in teaching artificial intelligence online. International Journal for the Scholarship of Technology-Enhanced Learning (1) 1

Goel AK, Joyner DA (2017) Using ai to teach ai: lessons from an online ai class. AI Magazine 38(2):48–59

Goel AK, Polepeddi L (2016) Jill watson: A virtual teaching assistant for online education. Tech. rep., Georgia Institute of Technology

Goel AK, Polepeddi L (2019) Jill watson: A virtual teaching assistant for online education. In: Dede C, Richards J, Saxberg B (eds) Learning Engineering for online education: Theoretical contexts and design-based examples., Routledge, chap 7

Gómez-Zará D, Paras M, Twyman M, Lane JN, DeChurch LA, Contractor NS (2019) Who would you like to work with? In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–15

Granovetter MS (1973) The strength of weak ties. American journal of sociology 78(6):1360–1380

Hostetter C, Busch M (2006) Measuring up online: The relationship between social presence and student learning satisfaction. Journal of the Scholarship of Teaching and Learning pp 1–12

Ischen C, Araujo T, Voorveld H, van Noort G, Smit E (2019) Privacy concerns in chatbot interactions. In: International Workshop on Chatbot Research and Design, Springer, pp 34–48

Jhaver S, Birman I, Gilbert E, Bruckman A (2019) Human-machine collaboration for content regulation: The case of reddit automoderator. ACM Transactions on Computer-Human Interaction (TOCHI) 26(5):1–35

Joyner DA, Isbell C, Starner T, Goel A (2019) Five years of graduate cs education online and at scale. In: Proceedings of the ACM Conference on Global Computing Education, pp 16–22

Kehrwald B (2008) Understanding social presence in text-based online learning environments. Distance Education 29(1):89–106

Kollock P (1997) Design principles for online communities

Kraut R, Egido C, Galegher J (1988) Patterns of contact and communication in scientific research collaboration. In: Proceedings of the 1988 ACM conference on Computer-supported cooperative work, pp 1–12

Kreijns K, Kirschner PA, Jochems W (2003) Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. Computers in human behavior 19(3):335–353

Larreamendy-Joerns J, Leinhardt G (2006) Going the distance with online education. Review of educational research 76(4):567–605

Lave J, Wenger E, et al. (1991) Situated learning: Legitimate peripheral participation. Cambridge university press

Lee M, Frank L, Beute F, De Kort Y, IJsselsteijn W (2017) Bots mind the social-technical gap. In: Proceedings of 15th European conference on computer-supported cooperative work-exploratory papers, European Society for Socially Embedded Technologies (EUSSET)

Lee YC, Yamashita N, Huang Y (2020) Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. Proceedings of the ACM on Human-Computer Interaction 4(CSCW1):1–27

Liao QV, Gruen D, Miller S (2020) Questioning the ai: Informing design practices for explainable ai user experiences. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp 1–15

Mayer JM, Jones Q, Hiltz SR (2015) Identifying opportunities for valuable encounters: Toward context-aware social matching systems. ACM Transactions on Information Systems (TOIS) 34(1):1–32

Motoyama M, Varghese G (2009) I seek you: searching and matching individuals in social networks. In: Proceedings of the eleventh international workshop on Web information and data management, pp 67–75

Narain J, Quach T, Davey M, Park HW, Breazeal C, Picard R (2020) Promoting wellbeing with sunny, a chatbot that facilitates positive messages within social groups. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp 1–8

Nordberg OE, Wake JD, Nordby ES, Flobak E, Nordgreen T, Mukhiya SK, Guribye F (2019) Designing chatbots for guiding online peer support conversations for

adults with adhd. In: International Workshop on Chatbot Research and Design, Springer, pp 113–126
Olsson T, Huhtamäki J, Kärkkäinen H (2020) Directions for professional social matching systems. Communications of the ACM 63(2):60–69
Ou C, Joyner DA, Goel AK (2019) Designing and developing video lessons for online learning: A seven-principle model. Online Learning 23(2):82–104
Patikorn T, Heffernan NT (2020) Effectiveness of crowd-sourcing on-demand assistance from teachers in online learning platforms. In: Proceedings of the Seventh ACM Conference on Learning@ Scale, pp 115–124
Ring L, Shi L, Totzke K, Bickmore T (2015) Social support agents for older adults: longitudinal affective computing in the home. Journal on Multimodal User Interfaces 9(1):79–88
Rohloff T, Sauer D, Meinel C (2020) Students' achievement of personalized learning objectives in moocs. In: Proceedings of the Seventh ACM Conference on Learning@ Scale, pp 147–156
Rovai AP (2001) Building classroom community at a distance: A case study. Educational technology research and development 49(4):33
Rovai AP (2002) Development of an instrument to measure classroom community. The Internet and higher education 5(3):197–211
Saha K, De Choudhury M (2017) Modeling stress with social media around incidents of gun violence on college campuses. Proceedings of the ACM on Human-Computer Interaction 1(CSCW):1–27
Simpson J, Gaiser F, Macík M, Breßgott T (2020) Daisy: A friendly conversational agent for older adults. In: Proceedings of the 2nd Conference on Conversational User Interfaces, pp 1–3
Skowron M, Tkalčič M, Ferwerda B, Schedl M (2016) Fusing social media cues: personality prediction from twitter and instagram. In: Proceedings of the 25th international conference companion on world wide web, pp 107–108
Slade S, Prinsloo P (2013) Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist 57(10):1510–1529
Sun N, Wang X, Rosson MB (2019) How do distance learners connect? In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp 1–12
Sundar SS (2020) Rise of machine agency: A framework for studying the psychology of human–ai interaction (haii). Journal of Computer-Mediated Communication 25(1):74–88
Szostek AM, Karapanos E, Eggen B, Holenderski M (2008) Understanding the implications of social translucence for systems supporting communication at work. In: Proceedings of the 2008 ACM conference on Computer supported cooperative work, pp 649–658
Terveen L, McDonald DW (2005) Social matching: A framework and research agenda. ACM transactions on computer-human interaction (TOCHI) 12(3):401–434
Wang Q, Camacho I, Goel A (2020a) Designing for agent-mediated online social connections: Lessons learned and potential challenges. URL

http://speechinteraction.org/CSCW2020/papers/3-Wang.pdf, presented at the Conversational User Interface Workshop at Computer Supported Cooperative Work (CSCW)

Wang Q, Jing S, Camacho I, Joyner D, Goel A (2020b) Jill watson sa: Design and evaluation of a virtual agent to build communities among online learners. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp 1–8

Wang Q, Jing S, Joyner D, Wilcox L, Li H, Plötz T, Disalvo B (2020c) Sensing affect to empower students: Learner perspectives on affect-sensitive technology in large educational contexts. In: Proceedings of the Seventh ACM Conference on Learning@ Scale, pp 63–76

Yorita A, Egerton S, Chan C, Kubota N (2020) Chatbot for peer support realization based on mutual care. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, pp 1601–1606

Zytko D, DeVreugd L (2019) Designing a social matching system to connect academic researchers with local community collaborators. Proceedings of the ACM on Human-Computer Interaction 3(GROUP):1–15

Zytko D, Regalado V, Grandhi SA, Jones Q (2018) Supporting online dating decisions with a prompted discussion interface. In: Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp 353–356