# Sonju: A Context-Aware Teachable Agent for Eliciting Older Adults' Situated Knowledge

JinYoung Yoo
School of Computing
KAIST
Daejeon, Republic of Korea
jinyoung.yoo@kaist.ac.kr

Somin Park
School of Computing
KAIST
Daejeon, Republic of Korea
som2n@kaist.ac.kr

Sueun Jang
School of Computing
KAIST
Daejeon, Republic of Korea
sueun.jang@kaist.ac.kr

Seungwan Jin*
Information & Electronics Research Institute
KAIST
Daejeon, Republic of Korea
seungwan@kaist.ac.kr

Uichin Lee*
School of Computing
KAIST
Daejeon, Republic of Korea
uclee@kaist.ac.kr

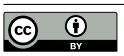
## Abstract

Older adults often face diminishing opportunities to express generativity. To address this gap, we propose **Sonju**, a teachable agent(TA) that supports older adults in teaching the knowledge and wisdom they have accumulated throughout their lives. Sonju is designed to generate context-aware questions using a large language model (LLM), store the knowledge conveyed by users in structured memory slots, and reuse it in later interactions to demonstrate continued learning and growth. As an initial feasibility study, we first examined whether activity contexts relevant to older adults could be recognized from mobile and wearable sensing data using the ExtraSensory dataset. We then evaluated the Conversation Layer in an end-to-end sensing-to-dialogue setting through simulated interactions with older-adult avatars exhibiting different levels of self-disclosure. For the end-to-end evaluation, we constructed a personalized activity-recognition model using sensing data from one participant. The model achieved 89.0% accuracy on a held-out test set, and its predicted activity contexts were provided to the Conversation Layer to initiate situated dialogue. The resulting dialogues and extracted knowledge were assessed using an LLM-based judge, providing preliminary evidence for the feasibility of the proposed approach. The evaluation suggested that Sonju maintained comparable conversational quality regardless of self-disclosure levels.

## CCS Concepts

• **Human-centered computing** → **Accessibility design and evaluation methods**; **Ubiquitous and mobile computing systems and tools**.

*Both authors co-advised this project.



This work is licensed under a Creative Commons Attribution 4.0 International License.
*UbiComp Companion '26, Shanghai, China*
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2533-3/2026/10
https://doi.org/10.1145/3798063.3837319

## Keywords

Context-Aware Teachable Agent; Older Adults; Generativity

**ACM Reference Format:**
JinYoung Yoo, Somin Park, Sueun Jang, Seungwan Jin, and Uichin Lee. 2026. Sonju: A Context-Aware Teachable Agent for Eliciting Older Adults' Situated Knowledge. In *Companion of the 2026 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '26), October 11–15, 2026, Shanghai, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3798063.3837319

## 1 Introduction

Older adults are not only recipients of care and assistance, but also active agents who can pass on the knowledge and experience they have accumulated over many years. In particular, teaching can be an important way to support older adults' generativity [4]. To support this, previous research has primarily focused on face-to-face interactions between older adults and recipients of their knowledge [10, 17]. While these interactions can foster real relationships and immediate exchange, they are also constrained by time, distance, and intergenerational hierarchy. [9]. To mitigate these issues, the teachable agent(TA) approach has been proposed, enabling older adults to experience a sense of self-efficacy by teaching a less knowledgeable agent [1, 3]. Much of the important knowledge older adults pass on is tacit, such as skills accumulated over time and memories associated with particular activities. Therefore, an agent designed to support older adults' teaching should be able to elicit and accumulate such knowledge within its relevant context [2]. We propose Sonju, a context-aware TA designed to provide older adults with opportunities to express generativity through teaching. Sonju consists of a ***Context Detection Layer*** for inferring everyday activities from sensing data and a ***Conversation Layer*** that uses an large language model(LLM) to generate questions suited to the given context. Older adults can respond by sharing their knowledge and experiences, which Sonju stores in structured memory slots and reuses in later interactions. By referring to previously learned knowledge, Sonju is designed to demonstrate continued learning and support older adults' sense of self-efficacy and connection with the agent. As an initial-stage study, we conducted two evaluations. First, we examined the feasibility of recognizing activity contexts relevant to

older-adult avators using the ExtraSensory Dataset [15]. Second, we evaluated the Conversation Layer through end-to-end sensing-to-dialogue interactions with older adult avatars exhibiting different levels of self-disclosure. For this evaluation, we constructed a personalized activity-recognition model using sensing data from one participant, which achieved an overall accuracy of 89.0% on a held-out test set. We also evaluated Sonju with an LLM-based judge using avatars with varying levels of self-disclosure, which suggested that Sonju elicited conversations comparably across these levels.

## 2 Design Requirements

To guide the design of a TA that supports older adults' generativity, we derived the following three design requirements.

### 2.1 DR1. Context-aware knowledge elicitation

The TA should infer the older adult's current activity context and use it to generate situated questions. Tacit knowledge is often difficult to articulate because it is embedded in everyday activities, physical settings, and personal experiences [11, 18]. Broad or decontextualized questions may therefore require older adults to recall and explain their knowledge without sufficient situational cues. For example, rather than asking about cooking in general, an agent may elicit richer experiential knowledge by asking about the dish being prepared or the current stage of preparation. Such contextual prompts can help older adults describe practical judgments, routines, and memories that may otherwise remain implicit. The agent should therefore use activity context inferred from sensing data to guide both the timing and content of its questions.

### 2.2 DR2. Active learner behavior

The agent should behave as an engaged learner rather than simply recording the user's utterances. Older adults may elaborate on their experiences when the learner acknowledges their responses and shows interest in what they are teaching [14]. A passive agent that merely collects statements may fail to communicate whether the teaching has been understood or valued. The agent should therefore react to the user's answers, ask follow-up questions, and invite clarification when an explanation is ambiguous. It should also reframe the topic or introduce a different perspective when the conversation stalls or when further elaboration may reveal additional experiential knowledge.

### 2.3 DR3. Visible agent growth

The agent should make its learning from the older adult visible over time. Remembering and reusing previous information can create continuity and rapport across interactions and help users recognize that their contributions have meaningfully shaped the agent [3, 5]. Simply storing information is not sufficient if the user cannot perceive that the stored knowledge affects later interactions. The agent should therefore refer to relevant prior teachings, connect new information with previously learned knowledge, and avoid asking questions that have already been answered. By demonstrating that it remembers and builds on what the older adult has taught, the agent can present itself as a learner whose knowledge develops through continued interaction.

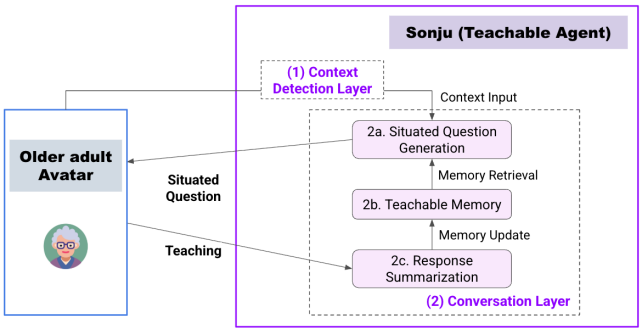



**Figure 1: Overview of the Sonju system architecture.**

## 3 System Design

We designed Sonju, an LLM-based TA that supports older adults in sharing the knowledge and experiences they have accumulated over time. The name Sonju means "grandchild" in Korean. Sonju takes the persona of a grandchild and interacts with older adults as a learner.

As illustrated in Figure 1, Sonju consists of two connected components: a ***Context Detection Layer*** and a ***Conversation Layer***. The Context Detection Layer infers everyday activity contexts, such as cooking, eating, and cleaning, from mobile and wearable sensing data and provides the detected activity as context input to the Conversation Layer. The Conversation Layer then uses this sensed context to initiate a situated interaction with the older adult.

The Conversation Layer consists of three main components. First, (2a) Situated Question Generation generates questions based on the current activity context and, when available, relevant knowledge retrieved from (2b) Teachable Memory. The situated question is presented to the older adult, who responds by teaching Sonju about their knowledge or experiences. The response is then processed by (2c) Response Summarization, which extracts and organizes the conveyed knowledge and updates Teachable Memory.

### 3.1 Context Detection Layer

Using the core sensor features from the ExtraSensory dataset [15], we reproduced an activity-recognition setting reported by Vaizman et al. [16]. With 177 features and 51 activity labels, the multi-task multilayer perceptron achieved a mean balanced accuracy (BA) [1] of 0.736. This was close to the BA of 0.739 reported for the corresponding model configuration, which used two hidden layers with 16 units each and no dropout.

We also examined six activity contexts selected for Sonju: cooking, eating, cleaning, doing laundry, watching television, and shopping. The mean BA across the six contexts was 0.706, with individual results ranging from approximately 0.65 to 0.73.

These results provide an initial assessment of the feasibility of recognizing everyday activity contexts relevant to Sonju from mobile and wearable sensing data. For the subsequent end-to-end evaluation of the Conversation Layer, we additionally constructed

[1] Balanced accuracy (BA) is the average of sensitivity and specificity: BA = (TPR + TNR)/2. It assigns equal importance to the positive and negative classes and is therefore less affected by class imbalance than conventional accuracy. We calculated balanced accuracy for each activity label and then averaged it across labels.

a personalized activity-recognition model using sensing data from one participant (P51) in the ExtraSensory dataset. This participant was selected for having the most complete label coverage across all seven target contexts, enabling training and test data construction for every context. Unlike the above feasibility analysis, the purpose of this model was to provide actual sensing outputs as inputs to Sonju's conversations. We retained all sensing instances, including periods in which none of the target activities occurred, as commonly observed in continuous everyday sensing. Each sensing instance was assigned to one of eight classes: cooking, eating, watching television, dressing, shopping, cleaning, doing laundry or none. We performed an 80/20 split separately within each class so that less frequent classes were represented in both the training and test sets. The resulting dataset consisted of 3548 training instances and 889 test instances. The activity contexts predicted by this personalized model were subsequently used as the sensed contexts in the end-to-end conversation evaluation described in 3.2.

### 3.2 Conversation Layer

The Conversation Layer was implemented as an LLM-based TA that takes on the role of a grandchild learning from an older adult. Each interaction comprises three phases: an opening that confirms the current activity, a main conversation that elicits activity-related knowledge and experiences, and a closing that summarizes what Sonju has learned. During the opening, Sonju generates a situated question based on the sensed activity context; if the context is incorrect, the older adult can correct it before continuing the conversation.

During the main conversation, the agent selects among three conversational strategies based on the older adult's utterances and the current dialogue state. The first strategy, follow-up questioning, elicits specific details about content introduced by the older adult. The second, reaction and confirmation, involves acknowledging or paraphrasing the response and checking whether the agent has understood it correctly. The third, perspective shifting, redirects the conversation when the current line of inquiry yields limited detail or when another perspective may elicit additional knowledge. These strategies are designed to provide more specific prompts when the older adult offers brief responses and to encourage further elaboration when richer responses are provided (DR2).

Sonju extracts knowledge conveyed in the older adult's utterances and stores it in structured memory slots. Each slot contains five fields: *context* (e.g., cooking), *sub-context* (e.g., cooking pasta), *category* (e.g., technical), *utterance* (e.g., "Taste it before it gets too soft"), and *note* (e.g., judging pasta doneness). The utterance field preserves the older adult's original statement, whereas the note provides a concise summary of the conveyed knowledge.

The category field includes three forms of generativity: technical generativity, which concerns practical skills and experiential know-how; cultural generativity, which concerns socially and historically acquired meanings and values; and parental generativity, which concerns practices and knowledge related to caring for and guiding younger family members [8, 13].

Stored memories persist across conversation sessions. At the beginning of a later session, the agent refers to previously acquired knowledge to demonstrate that it remembers the older adult's prior teachings. By retaining and building on this knowledge across sessions, Sonju makes its continued learning visible to the older adult (DR3).

## 4 Experiment & Results

### 4.1 Experiment Setup

Before conducting a user study with older adults, we conducted a simulation study to examine whether Sonju could elicit and retain knowledge from users with different levels of self-disclosure. The Conversation Layer was evaluated in an end-to-end sensing-to-dialogue setting, in which the activity contexts predicted by the personalized sensing model were used to initiate the simulated conversations.

We implemented an older-adult avatar as an LLM instance separate from Sonju, with the avatar taking the role of the teacher and Sonju taking the role of the learner. Each avatar was provided with activity-specific target knowledge and generated responses according to its assigned persona and knowledge base. The target knowledge included practical skills, family-related practices, and cultural meanings associated with each activity. The avatar was instructed to disclose this knowledge naturally in response to Sonju's questions rather than present it as a predefined list.

Because older adults vary in their levels of self-disclosure [6], and such variation may affect conversational dynamics [12], we created avatars with three self-disclosure profiles. In the *reserved* condition, the avatar provided brief responses containing only minimally necessary information. In the *moderate* condition, it provided an intermediate level of detail and occasionally elaborated when prompted. In the *talkative* condition, it proactively expanded on relevant experiences and knowledge.

**Table 1: Contexts and sub-contexts used in the elderly-avatar simulation**

| Context | Sub-contexts |
|---|---|
| Cooking | Cooking rice; Making kimchi; Cooking pasta |
| Eating | Eating dinner with family |
| Watching TV | Watching economic news; Watching a sports broadcast; Watching a history documentary |
| Doing laundry | Hand-washing clothes; Removing stains |
| Dressing | Wearing hanbok (traditional Korean clothes); Tying a necktie; Choosing clothes for the season |
| Shopping | Shopping at a traditional market; Choosing meat at a butcher shop |
| Cleaning | Removing mold |

The same set of 15 sub-contexts was used across all three persona conditions to avoid confounding topic variation with persona effects. The contexts and sub-contexts used to construct the target knowledge of the simulated older-adult avatars are summarized in Table 1. All interactions between Sonju and the simulated older-adult avatars were conducted in Korean.

The experiment comprised 45 conversation sessions: three persona conditions, each consisting of 15 sub-context sessions. At the beginning of each persona condition, Sonju's memory was fully reset to simulate interaction with a different older adult. Within

**Table 2: Evaluation Results across Older Adult Avatar Personas**

| Evaluation Dimension | Metric | Reserved | Moderate | Talkative |
|---|---|---|---|---|
| Conversation Quality | Naturalness | 96.5% | 95.7% | 95.7% |
| | Personalization | 100% | 100% | 91.7% |
| Generated QA Quality | Specific | 80% | 100% | 80% |
| | Inferred | 70% | 100% | 100% |
| Knowledge Extraction Quality | Correctness | 100% | 100% | 100% |
| | Coverage | 85.3% | 93.8% | 94.2% |
| | Completeness | 94.9% | 97.5% | 99.4% |

each condition, however, the memory was retained across all 15 sessions, allowing Sonju to refer to previously acquired knowledge and exhibit cumulative learning. Both Sonju and the older-adult avatar were implemented using GPT-4o, but they operated as separate instances with different role-specific prompts and independent conversation states.

## 4.2 Evaluation Results

*4.2.1 Personalized Context Detection for End-to-End Interaction.* For the end-to-end conversation evaluation, the personalized activity-recognition model achieved an overall accuracy of 89.0% on the held-out test set of 889 sensing instances. Among the context predictions used to initiate the end-to-end conversation sessions, 11 matched the corresponding true activity context, whereas 4 were incorrect detections. In these four cases, Sonju initially generated an opening question based on the incorrectly sensed context. The older-adult avatar corrected Sonju during the opening, after which the agent continued the conversation based on the true activity context.

*4.2.2 Conversation Evaluation.* We evaluated the Conversation Layer along three dimensions using an LLM-based judge [7], implemented with Claude Opus 4.8.

First, conversation quality was assessed using naturalness and personalization. Naturalness measured whether the agent's responses and questions were appropriate to the conversational context and whether the dialogue progressed coherently. Personalization measured whether the agent adapted its responses based on information from prior conversations. Both naturalness and personalization scores were similar across the three persona conditions, suggesting that Sonju maintained comparable conversational quality regardless of self-disclosure conditions. Second, generated-question quality was assessed for two question types. Specific questions explicitly referred to information from prior conversations, whereas inferred questions relied on plausible implications of the accumulated dialogue and memory. The questions were extracted using GPT-5 mini. Specific questions achieved appropriateness rates at least 80% across all conditions. Inferred questions reached 100% in the moderate and talkative conditions and 70% in the reserved condition. Third, knowledge extraction quality was evaluated using three metrics: correctness, coverage, and completeness. Correctness measured whether the extracted slots were grounded in the conversation. Coverage measured the proportion of the avatar's total knowledge that was extracted, whereas completeness measured the proportion of knowledge expressed during the conversation that was successfully extracted. Coverage increased with the avatar's level of self-disclosure, but remained 85.3% in the reserved condition. Completeness was at least 94.9% across all conditions, suggesting that Sonju captured most of the knowledge expressed during the simulated conversations.

## 5 Limitation & Future Work

The feasibility analysis of context detection was conducted using the ExtraSensory Dataset rather than sensing data collected directly from older adults. Although the results demonstrate the feasibility of recognizing activity contexts relevant to Sonju, they do not establish whether comparable recognition performance can be achieved for Sonju's intended population in everyday settings. In addition, this study provides an initial feasibility evaluation of whether the proposed TA can elicit and store tacit knowledge during simulated interactions with older-adult avatars. Although the current experiment directly connected sensing predictions to conversational interactions and included cases involving incorrect sensing outputs, the conversations were still simulated using LLM-based older-adult avatars. Future work should collect longitudinal sensing data from older adults and investigate personalized context-recognition models that can adapt to individual activity patterns in real-world environments. Furthermore, future studies should evaluate whether interactions with Sonju can enhance older adults' experience of generativity.

## 6 Conclusion

We proposed Sonju, a context-aware TA designed to support older adults' generativity through teaching. Sonju is designed to use activity context to generate situated questions, store the knowledge conveyed during interaction in structured memory slots, and reuse accumulated knowledge to make its continued learning visible. To evaluate the feasibility of Sonju's context recognition and its integration with conversation, we used the ExtraSensory Dataset to evaluate the system in an end-to-end sensing-to-dialogue setting, in which activity contexts predicted from sensing data were directly used to initiate Sonju's conversations. A personalized activity-recognition model constructed using data from one participant achieved 89.0% accuracy on a held-out test set. To evaluate the quality of the conversations while reflecting the characteristics of older adults, we designed three avatar profiles based on their levels of self-disclosure. An LLM-based evaluation confirmed that Sonju effectively elicited responses across all three avatar conditions. This study was limited to public sensing data and LLM-based simulation without actual older adults. Future work should evaluate the integrated sensing-to-dialogue system with sensing data and interactions collected from older adults and examine whether Sonju supports their expression of generativity and their experience of teaching a growing agent.

## Acknowledgments

Seungwan Jin and Uichin Lee co-advised this project. This work was supported by the 2026 AI Science Hub Program at the Korea Advanced Institute of Science and Technology (KAIST), funded by the Ministry of Science and ICT of the Republic of Korea (N10260120).

## References

[1] Ailiya Borjigin, Kaiyu Feng, and Zhiqi Shen. 2015. Designing Intergenerational Private Social Network with Teachable Agent. *International Journal of Information Technology* 21, 1 (2015).

[2] Anne Burmeister and Jürgen Deller. 2016. Knowledge retention from older and retiring workers: what do we know, and where do we go from here? *Work, Aging and Retirement* 2, 2 (2016), 87–104.

[3] Ping Chen, Xinjia Yu, Su Fang Lim, and Zhiqi Shen. 2024. Enhancing the well-being of seniors: A teachable agent for ikigai. *International Journal of Crowd Science* 8, 2 (2024), 71–76.

[4] Katie Ehlman and Mary Ligon. 2012. The application of a generativity model for older adults. *The International Journal of Aging and Human Development* 74, 4 (2012), 331–344.

[5] Agneta Gulz, Magnus Haake, and Annika Silvervarg. 2011. Extending a teachable agent with a social conversation module–effects on student experiences and learning. In *International conference on artificial intelligence in education*. Springer, 106–114.

[6] Yijie Guo, Ruhan Wang, Zhenhan Huang, Tongtong Jin, Xiwen Yao, Yuan-Ling Feng, Weiwei Zhang, Yuan Yao, and Haipeng Mi. 2025. Exploring the Design of LLM-based Agent in Enhancing Self-disclosure Among the Older Adults. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17.

[7] Rahul Gupta and Stephen DH Hsu. 2026. Embedded AI Companion System on Edge Devices. *arXiv preprint arXiv:2601.08128* (2026).

[8] J.N. Kotre. 1996. *Outliving the Self: How We Live on in Future Generations*. W.W. Norton. https://books.google.co.kr/books?id=hUQ0ZgWfiIMC

[9] Mandy HM Lau. 2024. Residential age segregation: Evidence from a rapidly ageing Asian city. *Journal of Population Ageing* 17, 3 (2024), 557–577.

[10] Karl Pillemer, Julia Nolte, Leslie Schultz, Harry Yau, Charles R Henderson Jr, Marie Tillema Cope, and Barbara Baschiera. 2022. The benefits of intergenerational wisdom-sharing: A randomized controlled study. *International journal of environmental research and public health* 19, 7 (2022), 4010.

[11] Michael Polanyi. 2009. The tacit dimension. In *Knowledge in organisations*. Routledge, 135–146.

[12] S Zahra Razavi, Lenhart K Schubert, Benjamin Kane, Mohammad Rafayet Ali, Kimberly Van Orden, and Tianyi Ma. 2019. Dialogue design and management for multi-session casual conversation with older adults. *arXiv preprint arXiv:1901.06620* (2019).

[13] Angela Schoklitsch and Urs Baumann. 2011. Measuring generativity in older adults. *GeroPsych* (2011).

[14] Megumi Tabuchi and Asako Miura. 2016. Intergenerational interactions when transmitting wisdom from older to younger generations. *Educational Gerontology* 42, 8 (2016), 585–592.

[15] Yonatan Vaizman, Katherine Ellis, and Gert Lanckriet. 2017. Recognizing detailed human context in the wild from smartphones and smartwatches. *IEEE pervasive computing* 16, 4 (2017), 62–74.

[16] Yonatan Vaizman, Nadir Weibel, and Gert Lanckriet. 2018. Context recognition in-the-wild: Unified model for multi-modal sensors and multi-label classification. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–22.

[17] Zinuo Wu, Tian Gan, Liwan Peng, Ziying Wang, Lai Liang, and Yan Zheng. 2026. Effects of Generativity Interventions on the Mental Health of Community-Dwelling Healthy Older Adults: A Pilot Study of the "Be My Mentor" Program. *Journal of Community Health Nursing* 43, 2 (2026), 122–140.

[18] Udo Zander and Bruce Kogut. 1995. Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. *Organization science* 6, 1 (1995), 76–92.