Human Factors in Design, Engineering, and Computing, Vol. 159, 2024, 82–91
https://doi.org/10.54941/ahfe1005570

Engaging All Elderly Residents in Community Renewal: Designer Spotlight Interview Tool for LLM Building

Ruochen Hu[1], Zihui Chen[1], Mengshi Yang[2], and Qiong Wu[1]

[1]Academy of Arts & Design, Tsinghua University, Beijing, China
[2]Shanghai International School of Design & Innovation, Tongji University, Shanghai, China

ABSTRACT

The natural language processing capabilities of Large Language Models (LLMs) can significantly enhance designers' ability to quantify unstructured information and improve communication with users, which is particularly important in rapidly aging societies. As elderly individuals engage in community renewal projects, they often face comprehension and expression barriers due to differences in cultural backgrounds and cognitive abilities, which complicates the acquisition of tacit knowledge for designers. To address this, we developed an AI community renewal toolkit, CommUnity AI, utilizing the fine-tuned ChatGPT-4o model. This toolkit provides easy-to-understand feedback to older adults through the creation of visual and textual information cards, and its effectiveness was evaluated in our study. The experiment involved 24 older adults and 6 designers, divided into experimental and control groups, and three separate focus group interviews were conducted. Using the SERVQUAL model to analyze the results, we found that the elderly participants showed greater trust and acceptance of the toolkit compared to traditional interview methods. CommUnity AI provides high-quality feedback through language comprehension, data collection, and visual and textual feedback, effectively reducing communication time while considering the needs and comprehension abilities of the elderly. This study underscores the potential of LLMs in community co-design, offering theoretical and practical insights into how designers can collaboratively engage with elderly individuals, ultimately fostering more inclusive and friendly community environments.

Keywords: Large language models, Human-computer interaction, Co-design, Age-appropriate tools, User experience evaluation

INTRODUCTION

Currently, community renewal is shifting from traditional physical space renovation to more comprehensive social space governance. Henri Lefebvre's socio-spatial theory emphasizes that space is not merely a physical entity but also a product of social relations and production activities. The existing design and functional configuration of public spaces in communities often follow a top-down planning approach, resulting in a disconnect between community construction and the dynamic needs of residents (Lai, 2023). This issue highlights the importance of emphasizing active resident participation in community planning and implementation to enhance the adaptability of community systems.

According to the latest data from the National Bureau of Statistics of China, by the end of 2023, the number of people aged 60 and above reached 290 million, making older adults a major component of communities (Du, 2023). Older adults in the community often face reduced social adaptability due to educational disparities and cognitive decline (Morse et al., 2024). These challenges lead to differences in understanding community building and participation, making it difficult for older residents to effectively express their opinions in community governance (Zhong et al., 2020). It is also challenging for designers and community workers to identify and address the real needs of older residents, leaving these groups isolated in community building (Cudjoe et al., 2020).

The natural language processing (NLP) capabilities of Large Language Models (LLM) offer new possibilities for older adults to participate in focus interviews for community building (Veres, 2022). These capabilities allow for a deep analysis of the expression habits of elderly residents, enabling designers to communicate in a manner that is more accessible to the elderly, transforming conversations into a list of needs and improving communication efficiency (Dwyer et al., 2024). LLMs can also include older residents from diverse backgrounds in community building programs during the interview process, promoting community co-creation activities (Shi et al., 2023).

This study developed the CommUnity AI toolkit to evaluate LLM's effectiveness in engaging older residents in community participatory design. Using the SERVQUAL model, we collected feedback from older residents and designers through a comparative experiment. This analysis enhanced older residents' involvement in focus group interviews and aided designers in adjusting strategies to ensure design solutions meet residents' needs, improving the applicability of community spaces and services.

METHOD

Materials

The team developed the innovative AI-based CommUnity AI toolkit by leveraging LLMs and GAI to enhance design capabilities through a structured process. Initially, a workshop was organized for elderly residents in Changping District, Beijing, to identify design points for the toolkit. The team then studied the principles of the LLM model, dividing it into three stages to better integrate focus interviews. Finally, a prototype framework using the LLM was built and integrated into the toolkit.

The workshop preparation focused on three core themes—space, culture, and community connection—using hierarchical topics, keywords, and images to help elderly and non-design participants better understand and express design goals (Borgianni et al., 2020).

We optimized LLM integration by modularizing its operation into information reception, processing, and feedback alongside design phases of information collection, demand analysis, and proposal development.

1. **Information Reception**

Designers and elderly residents collected data on community needs, which the toolkit's LLM analyzed to identify key needs, emotional tendencies, and potential issues.

2. **Information Processing**

The toolkit processes data to generate design proposals, with LLM refining core needs and priorities, leveraging generative AI for accurate and relevant content.

3. **Information Feedback**

The toolkit converts processed text into visual representations using ChatGPT-4o, creating keyword-linked images to enhance user understanding through intuitive visuals.

The information input and output framework is depicted in Figure 1.

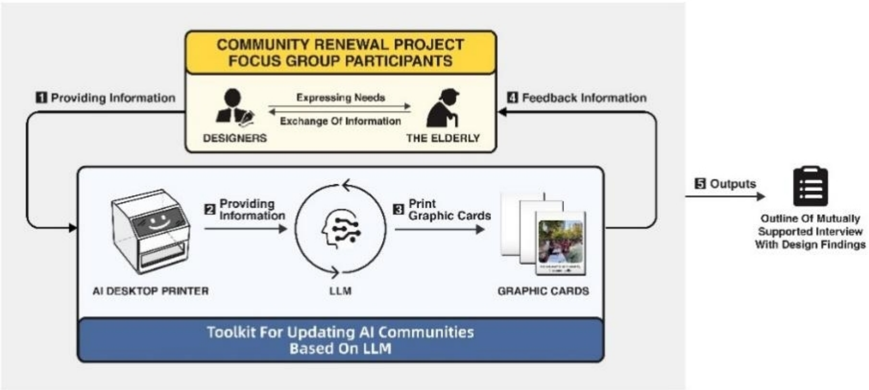

Figure 1: Information input and output framework diagram.

To enhance the LLM's effectiveness, we fine-tuned the GPT-4o model using extensive conversational data from older adults, including community interviews, questionnaires, and social media comments covering daily life, health needs, and social activities. The data was meticulously organized with contextual and emotional annotations, and model parameters were adjusted to include culturally appropriate greetings considering older adults' technology acceptance levels (Chen and Chan, 2014).

We integrated the trained LLM model into the desktop CommUnity AI device, which processes external information during focus interviews and produces illustrated cards for participants, as shown in Fig. 2.

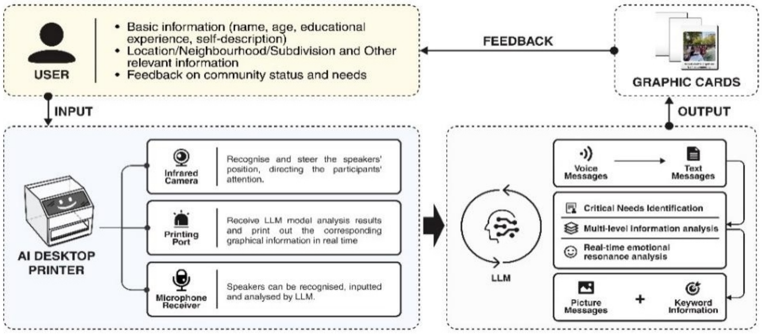

Figure 2: Community AI workflow.

Experiment Design

To verify the effectiveness of CommUnity AI, we conducted an experiment in a Beijing community involving 24 senior residents over 55 years old and 6 enterprise designers. Participant characteristics are detailed in Table 1.

Table 1. Information on participant characteristics.

	Community Resident											
	Gender		Age				Educational Level		Duration of Residence (Years)		Project Experience (Times)	
	Male	Female	55–60	61–65	66–70	Average Age	High School	Bachelor Degree+	5–7	10+	1–2	2+
Experimental group	5	7	4	6	2	61.6	6	6	5	7	9	3
Control group	5	7	4	5	3	62.1	7	5	5	7	8	4

	Gender		Project Experience		Interview Experience	
	Male	Female	3–5 Years	5+ Years	5–10 Times	10+ Times
Designer information	4	2	1	5	1	5

Each of the six focus group interviews involved two randomly selected designers conducting interviews with groups of four older adults. The control group used traditional interview methods, while the experimental group incorporated the CommUnity AI toolkit. We created an updated interview outline focusing on older adults' needs, preferences, abilities, motivations, and constraints to explore their dining needs in community cafeterias. Before the experiment, participants in the experimental group were instructed to familiarize themselves with the toolkit.

Data Collection

The SERVQUAL model proposed by Berry et al., in the late 1980s evaluates the "service quality gap" across five dimensions—tangibility, reliability, assurance, responsiveness, and empathy—based on the difference between actual service satisfaction and expectations (Shahin, 2003). To evaluate the tools used in our experiment (traditional interview tools and the CommUnity AI Toolkit), we administered a SERVQUAL-based questionnaire to participants in both the control and experimental groups. We used a Likert scale for quantitative assessment to ensure objectivity and comparability (Sullivan and Artino, 2013).

The Service Quality Evaluation Indicator System was constructed using relevant Chinese standards like GB/T 35796–2017 (see Table 2).

Table 2. CommUnity AI toolkit quality of service evaluation indicator system (Ruochen Hu et al., 2024).

Content of the Interviews		
Dimension	**Perspectives**	**Concern**
Tangibles	Environmental facilities	Q1. Display the layout of the environment according to the user's needs.
	Nutrition of dishes	Q2. Display detailed nutritional information and high-resolution pictures of the dishes.
	Service mode	Q3. Clearly display the service flow and operation guide.
	Service flow	Q4. Provide intuitive roadmap and navigation instructions with appropriate scenarios.
Reliability	Environmental facilities	Q5. Provide real-time updated facility status information.
	Nutrition of dishes	Q6. Ensure the accuracy and consistency of dish information.
	Service mode & Service flow	Q7. Low failure rate and high accuracy information processing capability.
Assurance	All	Q8. The content is expressed in a clear and concise manner and can be easily understood.
Responsiveness	All	Q9. Respond quickly to user inquiries and feedback and flexibly adjust services to adapt to changes in user needs.
Empathy	Environmental facilities	Q10. Quickly understand what the user is saying about a particular location and provide clear information.
	Nutrition of dishes	Q11. Customize the design of new models of food based on user feedback.
	Service mode	Q12. Provide suggestions for personalized services for older users.
	Service flow	Q13. Identify and assist in presenting the special needs of older users in the process.

After each focus group interview, elderly residents completed a SERVQUAL-based questionnaire on the spot, while the six designers completed their questionnaires after the two interviews they participated in. Cronbach's alpha coefficient analysis indicated strong internal reliability, with a Cronbach's alpha of 0.92 for satisfaction items (dimension scales ranging from 0.80 to 0.90) and 0.89 for expectation items (dimension scales ranging from 0.75 to 0.85), demonstrating good reliability and validity of the service quality evaluation questionnaire.

RESULTS & DISCUSSION

SERVQUAL Service Quality Assessment

In the control group, older community members' expectations for traditional tools averaged 4.050 ± 0.461, while their satisfaction averaged 3.320 ± 0.578, with a mean difference of 0.730 ± 0.672. For designers, the expectation mean was 3.970 ± 0.449, and satisfaction was 3.369 ± 0.552, resulting in a mean difference of 0.602 ± 0.669, as shown in Table 3.

Table 3. Control group data (Ruochen Hu et al., 2024).

SERVQUAL Dmensions for the Control Group						
	Community Resident			Designer		
Dimension	Mean Expectation Score (SE)	Mean Perception Score (SE)	Mean Gap Score (SE)	Mean Expectation Score (SE)	Mean Perception Score (SE)	Mean Gap Score (SE)
Tangibles (4)						
Q1	4.196±0.504	3.396±0.599	0.800±0.696	4.049±0.450	3.245±0.555	0.804±0.649
Q2	4.098±0.400	3.302±0.496	0.795±0.598	4.023±0.435	3.270±0.537	0.777±0.632
Q3	4.005±0.500	3.202±0.597	0.803±0.697	4.059±0.470	3.234±0.562	0.820±0.657
Q4	3.905±0.398	3.098±0.499	0.800±0.602	4.036±0.443	3.246±0.533	0.829±0.641
Reliability (3)						
Q5	3.998±0.502	3.600±0.600	0.397±0.696	3.901±0.448	3.496±0.553	0.405±0.654
Q6	3.901±0.397	3.505±0.502	0.404±0.603	3.909±0.449	3.499±0.554	0.410±0.655
Q7	3.800±0.504	3.395±0.598	0.401±0.698	3.903±0.447	3.494±0.552	0.409±0.653
Assurance (1)						
Q8	4.198±0.396	4.001±0.499	0.205±0.605	4.197±0.399	3.998±0.504	0.199±0.602
Responsiveness (1)						
Q9	3.699±0.502	2.903±0.605	0.802±0.799	3.702±0.502	2.900±0.601	0.802±0.797
Empathy (4)						
Q10	4.002±0.504	3.197±0.595	0.803±0.697	4.036±0.458	3.237±0.559	0.799±0.656
Q11	3.903±0.399	3.103±0.505	0.801±0.601	3.976±0.443	3.171±0.544	0.805±0.643
Q12	4.104±0.403	3.302±0.499	0.804±0.604	4.045±0.465	3.284±0.569	0.761±0.668
Q13	3.999±0.504	3.202±0.604	0.797±0.698	4.027±0.448	3.225±0.557	0.802±0.649

Descriptive statistics for the CommUnity AI Toolkit factors are shown in Table 4. The gap between the actual experience and expectations was minimal, with older participants having an average expectation score of 4.741 ± 0.351 and a satisfaction score of 4.550 ± 0.451, resulting in a difference of –0.191 ± 0.601. Designer feedback showed an expectation mean of 4.741 ± 0.351 and a satisfaction mean of 4.499 ± 0.451, with a difference of –0.242 ± 0.596, indicating better performance than the control group.

Table 4. Experimental group data (Ruochen Hu et al., 2024).

SERVQUAL Dimension of the Experimental Group						
	Community Resident			Designer		
Dimension	Mean Expectation Score (SE)	Mean Perception Score (SE)	Mean Gap Score (SE)	Mean Expectation Score (SE)	Mean Perception Score (SE)	Mean Gap Score (SE)
Tangibles (4)						
Q1	4.901±0.404	4.701±0.499	-0.202±0.604	4.762±0.354	4.446±0.453	-0.316±0.553

(Continued)

Table 4. Continued

SERVQUAL Dimension of the Experimental Group						
	Community Resident			Designer		
Dimension	Mean Expectation Score (SE)	Mean Perception Score (SE)	Mean Gap Score (SE)	Mean Expectation Score (SE)	Mean Perception Score (SE)	Mean Gap Score (SE)
Q2	4.802±0.300	4.599±0.398	-0.202±0.502	4.748±0.355	4.748±0.355	-0.308±0.552
Q3	4.703±0.403	4.502±0.500	-0.201±0.596	4.751±0.356	4.441±0.450	-0.310±0.554
Q4	4.599±0.300	4.403±0.400	-0.203±0.501	4.740±0.348	4.445±0.449	-0.295±0.548
Reliability (3)						
Q5	4.700±0.398	4.498±0.501	-0.198±0.596	4.602±0.351	4.366±0.449	-0.236±0.565
Q6	4.597±0.296	4.397±0.404	-0.197±0.503	4.598±0.349	4.364±0.447	-0.234±0.563
Q7	4.496±0.404	4.296±0.495	-0.199±0.596	4.600±0.350	4.365±0.448	-0.235±0.564
Assurance (1)						
Q8	4.398±0.300	4.201±0.400	-0.202±0.496	4.197±0.399	3.998±0.504	0.199±0.602
Responsiveness (1)						
Q9	4.798±0.400	4.499±0.602	-0.297±0.697	3.702±0.502	2.900±0.601	0.802±0.797
Empathy (4)						
Q10	4.698±0.404	4.401±0.503	-0.297±0.603	4.762±0.354	4.401±0.451	-0.361±0.553
Q11	4.600±0.304	4.304±0.398	-0.297±0.504	4.751±0.355	4.398±0.452	-0.353±0.554
Q12	4.904±0.304	4.501±0.403	-0.398±0.495	4.748±0.356	4.399±0.450	-0.349±0.553
Q13	4.801±0.400	4.400±0.500	-0.396±0.599	4.740±0.348	4.402±0.448	-0.338±0.551

Feedback for Elderly Participants

Based on Figure 3, we derived three key conclusions: enhanced engagement, real-time empathetic feedback, and effective AI-assisted community building.

Figure 3: Attitudes of the elderly towards CommUnity AI VS. Traditional interviewing tools.

Enhanced Engagement

Our study demonstrates the effectiveness of the CommUnity AI toolkit in enhancing older adults' engagement in community renewal programs. By generating personalized questions through a detailed community information repository, the toolkit led to higher satisfaction among older adults regarding information accuracy and consistency. This increased trust in the toolkit and encouraged more open expression of opinions, highlighting the shortcomings of existing tools in supporting community renewal initiatives.

Real-Time Empathetic Feedback

The toolkit's real-time feedback, graphic presentations, and adaptive features make older residents feel understood and valued, fostering empathy and confident participation in the community renewal process.

Effective AI-Assisted Community Building

The toolkit can creatively address diverse challenges during conversations, supporting the idea of AI-assisted community building proposed by Gubing W et al., This concept suggests that AI can inclusively engage a more diverse group of 'age-appropriate' older adults.

Feedback for Designers

Based on Figure 4, we derived two key conclusions: enhanced design process and better human-computer collaboration.

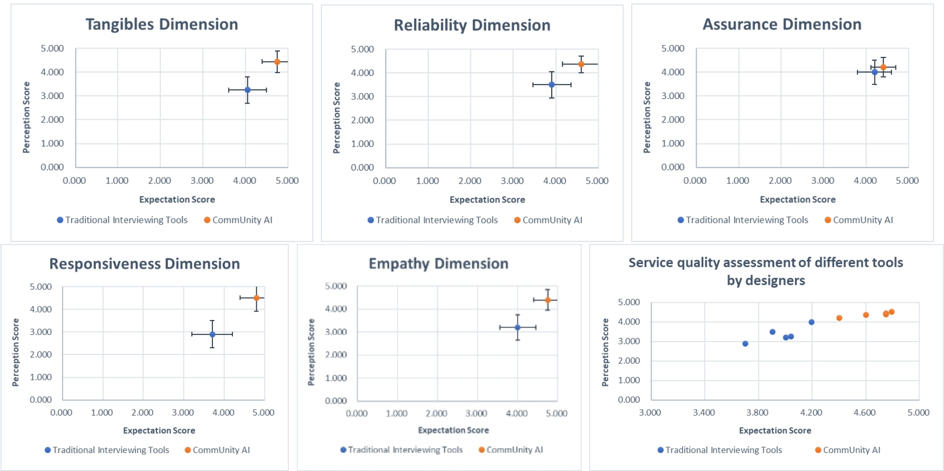

Figure 4: Attitudes of designers towards CommUnity AI VS. Traditional interviewing tools.

Enhanced Design Process

The CommUnity AI Toolkit exhibited greater rigor and a wealth of accumulated experience, effectively assisting designers in developing design solutions that better meet community needs. Its auto-generation and updating features reduced pressure on designers during interviews, minimizing human error and information omission, and accelerating the design process. The toolkit's high accuracy in graphical and data analysis helped create a comprehensive database on community infrastructure and resource allocation, offering deeper insights into older residents' behaviors. These findings boosted designers' confidence, leading to more relevant and effective design solutions.

Improved Human-Computer Collaboration

Additionally, the toolkit supports designers with data and preliminary analysis, facilitating efficient collaboration on community renewal projects. Its empathetic capabilities help designers create detailed user profiles, understand latent needs, and enhance design quality, increasing project success rates.

LIMITATONS

Research is needed to identify effective interaction patterns and improve the emotional warmth of AI communications for older adults, especially those with lower literacy and education levels.

For designers, to ensure a balanced relationship, it's crucial to strengthen collaboration and refine workflows between AI toolkits and designers, preventing the undervaluation of human experience.

CONCLUSION

In summary, utilizing CommUnity AI can effectively reduce communication barriers within the community and better address the needs of older adults through real-time feedback and visual information representation. This approach also provides designers with more comprehensive feedback and data. Our research identifies current issues, including the humanistic aspects of AI and the balance between designers and AI tools. To address these gaps, future studies should implement more pilot programs to evaluate these factors across different community settings.

The study's findings provide a foundation for developing design renewal strategies with CommUnity AI, enhancing community life quality and offering a valuable tool for designers.

ACKNOWLEDGMENT

This work is supported by the "Dual High" project of Tsinghua Humanity Development (No. 2021TSG08203).

REFERENCES

Borgianni, Y., Cascini, G., Rotini, F. and van Wodehouse, A. (2020) 'Forms of stimuli and their effects on idea generation in terms of creativity metrics and non-obviousness', International Journal of Design Creativity and Innovation, 8(3), pp. 147–164.

Chen, K. and Chan, A. H. S. (2014) 'Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM)', Ergonomics, 57(5), pp. 635–652.

Cudjoe, T. K. M., Roth, D. L., Szanton, S. L., Wolff, J. L., Boyd, C. M. and Thorpe, R. J. (2020) 'The epidemiology of social isolation: National health and aging trends study', Journal of Gerontology: Series B, Psychological Sciences and Social Sciences, 75(1), pp. 107–113.

Du, P. (2023) 'Current situation of the Chinese population aging and the development of social security system', Social Science Abstracts, 2023(7), pp. 8–10.

Dwyer, N., Harrison, M., O'Mara, B. and Ryan, M. (2024) 'Inclusive gaming through AI: a perspective for identifying opportunities and obstacles through co-design with people living with MND', Frontiers in Computer Science, 6.

Emerson, R. W. (2017) 'Likert scales', Journal of Visual Impairment & Blindness, 111(5), p. 488.

Lai, S. (2023) 'Function orientation, system adaptation, and governance mechanisms in community renewal from the perspective of social space theory', Social Security and Administration Management, 4(8).

Morse, M. R., Koutsoubelis, F., Whitfield, T. and Wells, C. (2024) 'Worry and ruminative brooding: associations with cognitive and physical health in older adults', Frontiers in Psychology, 15, pp. 1332398–1332398.

Shahin, A. (2003) 'SERVQUAL and model of service quality gaps: a framework for determining and prioritizing critical factors in delivering quality services', in Proceedings of the 10th International Conference on Quality. St. Petersburg, Russia: The Russian Association for Quality.

Shi, Y., Gao, T., Jiao, X. and Cao, N. (2023) 'Understanding design collaboration between designers and artificial intelligence: a systematic literature review', Proceedings of the ACM on Human-Computer Interaction, 7.

Sullivan, G. M. and Artino, A. R. (2013) 'Analyzing and interpreting data from Likert-type scales', Journal of Graduate Medical Education, 5(4), pp. 541–542.

Veres, C. (2022) 'Large language models are not models of natural language: they are corpus models', IEEE Access, 10, pp. 61970–61979.

Zhong, S., Lee, C., Foster, M. J. and Bian, J. (2020) 'Intergenerational communities: a systematic literature review of intergenerational interactions and older adults' health-related outcomes', Social Science & Medicine, 264, 113374.