Modular AI agents for transportation surveys and interviews: Advancing engagement, transparency, and cost efficiency

Jiangbo Yu[a,], Jinhua Zhao[b], Luis Miranda-Moreno[a], Matthew Korp[c]

[a] *Department of Civil Engineering, McGill University, Montreal, H3A 0C3, Canada*
[b] *Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, 02139, USA*
[c] *Department of Mechanical Engineering, McGill University, Montreal, H3A 0C3, Canada*

E-mail: jiangbo.yu@mcgill.ca

Surveys and interviews—structured, semi-structured, or unstructured—are widely used for collecting insights on emerging or hypothetical scenarios. Traditional human-led methods often face challenges related to cost, scalability, and consistency. For example, distributed questionnaires lack the ability to provide real-time guidance and request immediate clarifications. Recently, various domains have begun to explore the use of conversational agents (chatbots) powered by generative artificial intelligence (AI) technologies. However, considering decisions in transportation investments and policies often carry significant public and environmental stakes, surveys and interviews face unique challenges in integrating AI agents, underscoring the need for a rigorous, resource-efficient approach that enhances participant engagement and ensures privacy. This paper addresses this gap by introducing a modular approach and its resulting parameterized process for designing AI agents. We detail the system architecture, integrating engineered prompts, specialized knowledge bases, and customizable, goal-oriented conversational logic. We demonstrate the adaptability, generalizability, and efficacy of our modular approach through three empirical studies: (1) travel preference surveys, highlighting conditional logic and multimodal (voice, text, and image generation) capabilities; (2) public opinion elicitation on a newly constructed, novel infrastructure project, showcasing question customization and multilingual (English and French) capabilities; and (3) expert consultation about the impact of technologies on future transportation systems, highlighting real-time, clarification request capabilities for open-ended questions, resilience in handling erratic inputs, and efficient transcript postprocessing. The results suggest that the AI agent increases completion rates and response quality. Furthermore, the modular approach demonstrates controllability, flexibility, and robustness while addressing key ethical, privacy, security, and token consumption concerns. We believe this work lays the foundation for next-generation surveys and interviews in transportation research.

We present three concrete cases in which the proposed architecture and process flow for conducting S&Is are utilized. The user interfaces (dialog boxes) for all three experiments are accessible through smartphones and computer web browsers. The proposed modular approach is platform-agnostic. As we are writing the paper, various online platforms are available for AI agent development and deployment. To clarify the role of the proposed modular approach and platforms for AI agent development and deployment, we can draw an analogy to designing a bridge with computer-aided design (CAD) software/platform. Just as CAD software supports and facilitates an engineer in creating a complex structure (the bridge), a platform for AI agent development and deployment can help realize and implement our proposed modular approach in building an AI agent for S&Is. Criticizing the methodological contribution of our approach because it utilizes AI development platforms would be akin to attributing all credit for a bridge's design to the CAD software. The proposed approach is also module-agnostic. That is, any module shown in Fig. 1 is replaceable as long as the new module can serve the same function. For example, one can choose any open-source LLM for the on-premise LLM module. For another example, one can design their own user interface (UI module) using a preferred programming language.

The online LLMs used in the first, second, and third experiments are from GPT-4, GPT-4o, and Claude 3-SonNet, respectively. In addition to using Llama 3.2 for the on-premise LLM component for privacy protection, we implemented a set of additional measures to ensure the protection of participants' data. First, we make it optional for the participants to decide whether they would like to share their (potentially) identifiable information. Second, even if they share their (potentially) identifiable information such as name, workplace, and email address, we employed anonymized techniques to dissociate any personally identifiable information in data processing and in released results. For home location information, we only ask participants for the first three (of the six) digits of the postal code of their home address in Canada. The response goes through an on-premise LLM to detect any privacy concerns. Only then will the process manager module consider whether to send the information to cloud-based LLMs for further processing. The on-premise LLM might also suggest that the process manager hide or modify certain sensitive information before being sent to an online commercial LLM. Furthermore, we prioritized secure data storage practices. The collected data were stored on dedicated hardware that featured security protections provided by McGill University. Only authorized researchers and supervisors were granted access to this hardware, and stringent access control mechanisms were in place to safeguard against unauthorized access or breaches. IP addresses are not associated with usernames or user IDs, and they are only used to identify when a new request is needed during the operation of an AI agent, whether to continue from the previous chat (if any) or launch a completely new conversation.

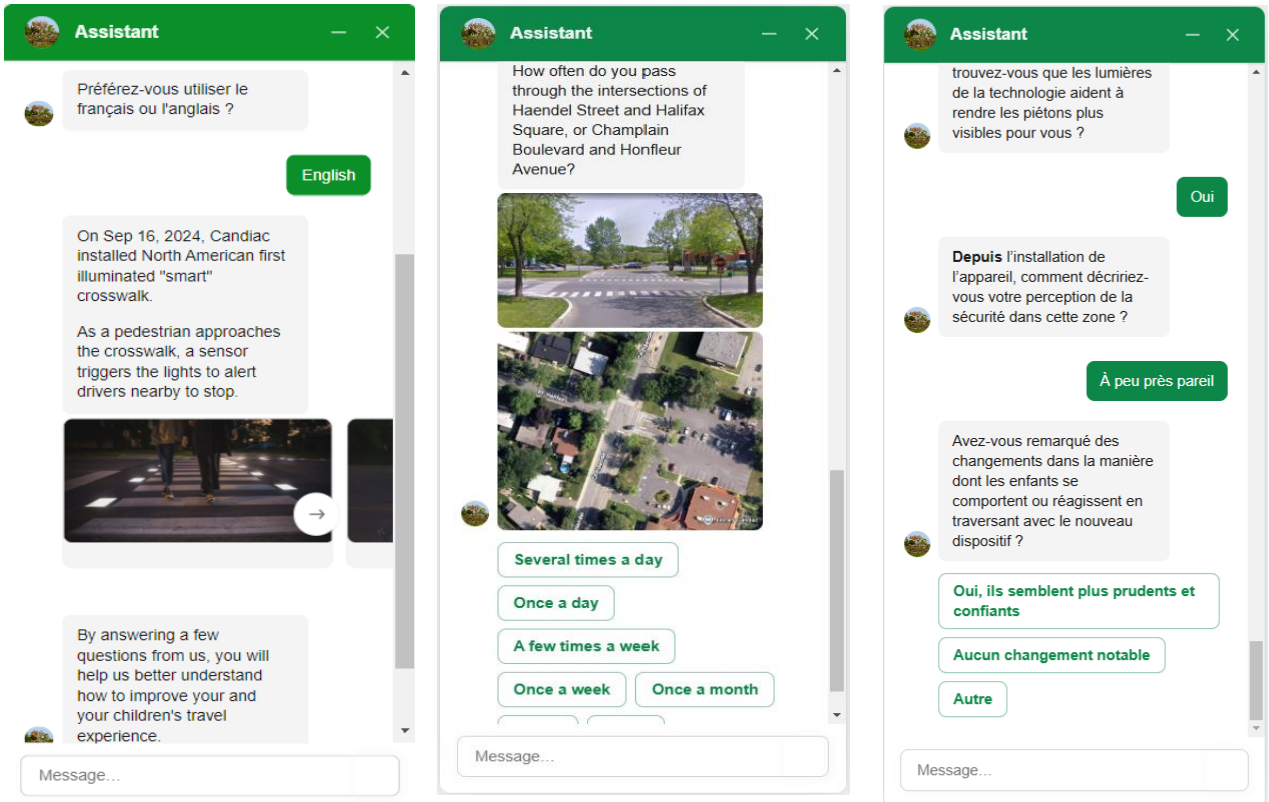
Approval for all ethical and experimental procedures and protocols was granted by the Research Ethics Board (REB) office at McGill University and by the Committee on the Use of Human Experimental Subjects (COUHES) at the Massachusetts Institute of Technology.

Experiment: Public perception for infrastructure project

Due to page limit, we only present the second study develops and deploys an AI agent using the proposed modular approach for understanding the public perception of a recently (September 2024) implemented pedestrian crosswalk lighting system in the City of Candiac, Quebec, Canada, which is considered the first of its kind in North America (CBC News, 2024). The lighting system is shown in Fig. 1. The targeted participants of the students at the St-Mark's Elementary School adjacent to the new lighting system.

Fig. 1 Newly installed lighting system in Candia, Quebec, Canada, for improving intersection traffic safety. The image is adopted with permission from the Ville de Candiac [Source].

The City of Candiac tested the AI agent and distributed it among residents (a major portion is the parents from the nearby elementary school), with an email containing a URL and a QR code to the agent. Participants can directly click the link on their smartphone or use their laptop or tableau to scan the QR code. The interface of the agent is shown in Fig. 2. For comparison, we also developed a web-based questionnaire version of the same set of questions. The difference is that the web-based version does not have AI assistance and the ability to be conversational, such as paraphrasing participants' responses to ensure clear understanding and asking for clarifications. The web-based form also does not allow voice-based responses. Our system randomly distributed 80% of the participants for the web-based questionnaire and the other 20% for the AI agent. In both cases, the survey starts by informing participants about the study's objective, the agent's functionalities, and the data privacy measures. The participants can choose to use French or English modes. Note that, in the case of AI agent, there are no two separate process flows for each language; instead, it is one single process flow, where the LLM can take French or English as inputs and produce output using the language specified in the associated prompt parameterized by a variable (about what language to use) based on the participant's response to the initial question about the preferred language. The agent also allows the participants to interact through voice commands, which was particularly beneficial for those who preferred speaking over typing. In discrete choice questions, the agent attempts to match the voice answer with the options available. If there is no match, the agent will ask for clarification. In open-ended questions, the agent attempts to summarize or paraphrase the participants' responses and asks for clarifications when necessary to ensure data accuracy. If the agent considers a response to be challenging to summarize or paraphrase (likely due to response irrelevance), the agent asks the participant to clarify with more detail and paraphrase the participant's response to ensure its understanding.



(a) (b) (c)

Fig. 2 Agent interface during an interview. The agent can use English or French to communicate with the participants and follow up for further clarifications.

The process consists of three main phases (each phase has approximately 7 questions): (1) collect information about the travel mode, frequency, and time of day for the participants to pass through the intersections or send their child (children) to the school and whether they have noticed the lighting technology; (2) based on the participants' responses, the agent will either ask whether the parents think the technology is effective in terms of safety improvement or explain the technology is about; and (3) elicit opinions (including both discrete choice and open-ended questions) about any other thoughts and suggestions and whether the parents think the technology should be installed in more intersections in the region.

A total of 117 participants started the survey with the agent, and 29 entered the web-based survey (without AI enhancement). We did not provide participation incentives in either formats except that we remind the participants that their responses will be helpful for the city to better understand how to improve their travel experience in the future. Of the participants who started the survey with the agent, 78 (approximately 66.7%) completed the survey with sufficient quality; Of the participants who entered the web-based survey, 11 (approximately 37.9%) completed the web-based survey with sufficient quality. The difference in survey completion percentages suggests that the conversational nature of the AI agent might promote more participation than the web-based form (not conclusive due to the relatively small sample size).

We further prompted the AI agent to only elicit the opinions but also postprocess the responses into a standard data analytics (spreadsheet) format. We manually checked the organized data and found a high level of accuracy. Since this section primarily aims to demonstrate the efficacy of the agent using the modular approach, and we have not yet obtained permission to disclose the full survey results, we present only preliminary findings in Fig. 3 for demonstration purposes. Fig. 3a illustrates that 85% of participants passed through the intersection at least once a month in the past year. Fig. 3b shows that approximately 74% have used the smart crosswalk, while Fig. 3c indicates that over 76% traversed the intersection on foot or by bike, though some may have also done so by car. Fig. 3d highlights that more than 62% frequently cross the intersection at night. Regarding perceptions of the new lighting system's impact on safety, Fig. 3e shows that 50% believe it enhances safety, while 40% do not. Among those who walked or biked through the intersection with the newly installed lighting system, Fig. 3f reveals that 39% found drivers' behavior toward them to be more considerate. Fig. 3g indicates that 88% of respondents believe the new lighting system improves nighttime visibility in regular weather conditions, and Fig. 3h shows that over half of the participants find that it enhances visibility in snowy conditions. Finally, Fig. 3i suggests that about 55% of respondents think similar lighting systems could improve safety in other locations. Overall, these preliminary findings suggest that most participants familiar with the newly installed crosswalk lighting system hold neutral to positive views about the project. However, as this lighting technology is still relatively new, it remains uncertain whether perceptions will shift with increased familiarity over time. Future studies are needed to assess the long-term stability of these initial trends, and we leave the comparison between responses obtained from the agent and those from the web-based questionnaire for future research.

The experiment demonstrated the effectiveness of the LLM-enabled chatbot in conducting public outreach for infrastructure projects. Approximately 20% of the participants chose to use voice, and 83% chose French, which suggests that the agent's voice capabilities and bilingual functions are valuable for encouraging S&I participants who otherwise would not participate. We also noticed that in approximately 4% of the discrete-choice questions and 30% of the open-ended questions, the agent asked for clarifications (we set the maximum number of clarification requests to two), and we did find that the clarification provided clearer and more definite answers.

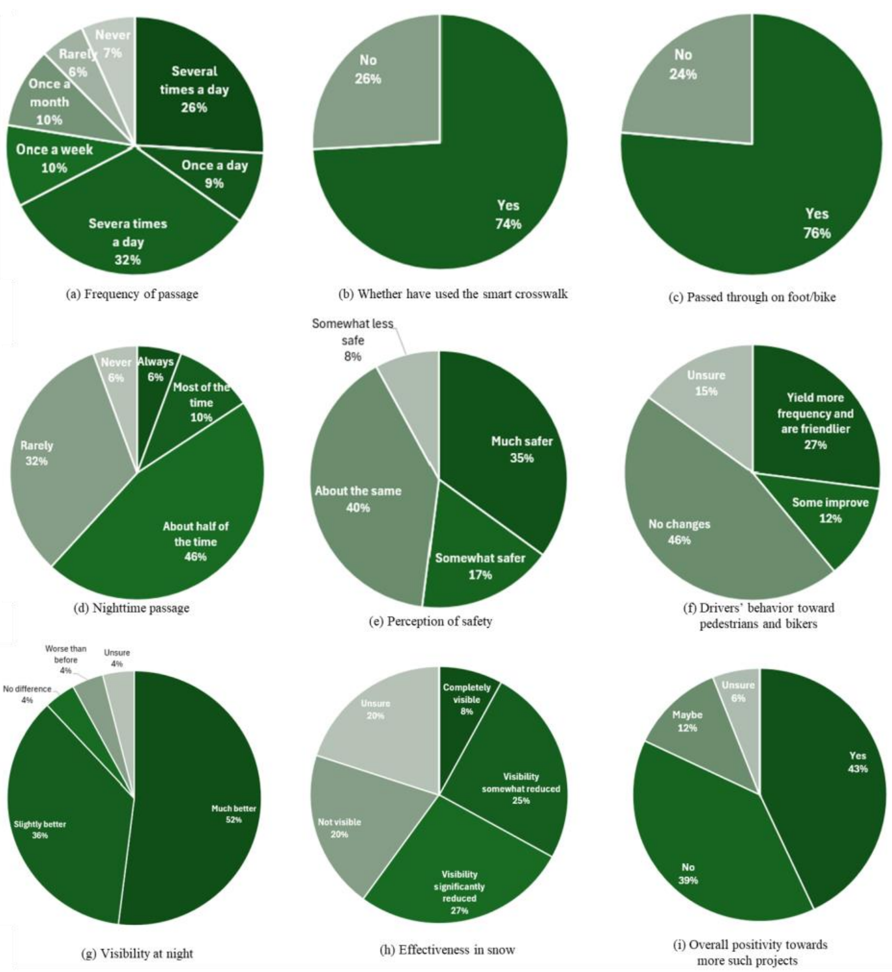


Fig. 3 Preliminary survey results (unweighted) on usages and perceptions of the smart crosswalk lighting system