# Ekai Hashimoto

Nagoya Institute of Technology
Nagoya, Aichi
Japan

e.hashimoto.611@stn.nitech.ac.jp
https://ekaigithub.github.io/mypage/

## 1 Research interests

My research interests are building a spoken dialogue system (SDS) that engages deeply with users' "sense of values," particularly in domains such as career interviews. To realize the system, I focus on two main areas: enhancing the information collecting abilities of slot-filling dialogue systems (SFD systems) using large language models (LLMs), and exploring the potential of SDS in supporting career development.

### 1.1 Improving collecting information ability

Slot-filling-based dialogue management, where a set of slots and utterances represents knowledge acquired during the conversation, is generated based on that slot set and has long been utilized in dialogue systems (Bobrow et al., 1977). However, since a predefined set of slots controls the dialogue, there have been limitations in information collecting ability, especially in adapting to career interview scenarios. To address this issue, Finch et al. (2024) adds new slots based on clustering of independent value spans. However, it still faces challenges such as missing important information and generating redundant slots. This approach lacks the flexibility to form schemas dynamically according to the dialogue's context and structure.

To overcome these limitations, we incorporated the abductive reasoning process prompt into LLM-based slot generation to extract deeper information from users (Hashimoto et al., 2025). This allows the system to collect contextually rich information from users facing major career decisions, such as resignation or reassignment, while still grounded in a predefined initial slot set.

### 1.2 Questions Ordering

In career interviews for nurses considering leaving their jobs, asking "Do you want to quit?" at the outset rarely elicits an honest answer. Trust-building has traditionally relied on rapport-establishing techniques such as empathic responses and small talk to address this issue. (Bickmore and Cassell, 2001) However, interactions between humans and systems are influenced by multiple factors, beyond interpersonal cues, including anonymity, power dynamics, and memory retention. Therefore, simply adjusting the tone of a system's utterances is unlikely to encourage self-disclosure sufficiently.

This study categorizes questions into four stages based on their level of "depth":
1. low-load factual confirmation
2. assessment of values and satisfaction
3. recognition of future vision and current gaps
4. exploration of core emotions and turnover intention

Next, we investigate which ordering these levels naturally encourages users to self-disclose. To ensure, we use the Japanese counselling dataset (about 6k long-form dialogues)(Qi et al., 2025) and the Emotional-Support-Conversation corpus (about 10k sessions)(Liu et al., 2021). We analyse question depth and topic transitions quantitatively, then embed the resulting patterns as state transition graphs in dialogue-state tracking (DST) to evaluate their effect on improving turnover intent estimation.

## 2 Spoken dialogue system (SDS) research

**Dialogue research in 5 to 10 years** Recent advances in LLMs have dramatically boosted the performance of SDSs. Over the next five to ten years, SDSs will incorporate long-term conversational history and psychological insights, enabling far richer user understanding and more human-like interaction. As such systems become widely deployed, the resulting accumulation of dialogue data and user feedback should create a virtuous cycle that further accelerates performance gains.

At the same time, the very need for dialogue is beginning to waver. Thanks to LLMs, benefits that once required conversational exchange, information, reassurance, and a sense of conviction can increasingly be delivered through non-dialogue channels. If this trend continues, dialogue may be dismissed as costly, risking weaker social bonds and more isolated decision-making. For this reason, future dialogue research must confront the fundamental questions: Why do we converse? What distinctive value does dialogue offer? And how can we protect and amplify that value in an age where non-dialogue alternatives proliferate?

**Societal Deployment and Evaluation of SDSs** Traditional dialogue-system research has focused on technical completeness, concentrating on goal-oriented dia-

logues and evaluating performance with quantitative metrics such as slot-filling accuracy and intent matching. With the rise of large language models (LLMs) that let almost anyone build high-functioning systems, there is an increasing need for new evaluation frameworks prioritizing user experience in real-world scenarios. The dialogue system we are developing serves as a component for collecting information from nurses during career-counseling interviews; we therefore place particular emphasis on system design and evaluation within a broader social context. Therefore, our works also emphasize system design and evaluation in a broader social context.

Based on this perspective, we believe that dialogue systems require multi-layer evaluation, not only from the standpoint of end users (e.g., nurses) but also from other stakeholders such as system deployers (e.g., hospitals) and system designers (e.g., dialogue-system researchers). Although it may not be necessary to assess all of these layers at once, considering them can help prevent structural misalignments, for example, optimizing for short-term user satisfaction at the expense of longer-term impacts on other layers.

## 3 Suggested topics for discussion

I would like to discuss the following topics:

- Difficulties in evaluating and setting experiment dialogue systems aimed at promoting users' internal growth and long-term use

- Strategies for cultivating interdisciplinary partnerships (psychology, conversation analysis) to validate novel dialogue-management methods empirically

- In the fields where AI is not expected to replace human decision-making: To what extent should AI systems exercise autonomy? How should they participate in decision processes led by humans?

### Acknowledgments

### References

Timothy Bickmore and Justine Cassell. 2001. Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, CHI '01, page 396–403. https://doi.org/10.1145/365024.365304.

Daniel G. Bobrow, Ronald M. Kaplan, Martin Kay, Donald A. Norman, Henry Thompson, and Terry Winograd. 1977. GUS, a frame-driven dialog system. *Artificial Intelligence* 8(2):155–173. https://doi.org/10.1016/0004-3702(77)90018-2.

James D. Finch, Boxin Zhao, and Jinho D. Choi. 2024. Transforming slot schema induction with generative dialogue state inference. In Tatsuya Kawahara, Vera Demberg, Stefan Ultes, Koji Inoue, Shikib Mehri, David Howcroft, and Kazunori Komatani, editors, *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Association for Computational Linguistics, Kyoto, Japan, pages 317–324. https://doi.org/10.18653/v1/2024.sigdial-1.27.

Ekai Hashimoto, Mikio Nakano, Takayoshi Sakurai, Shun Shiramatsu, Toshitake Komazaki, and Shiho Tsuchiya. 2025. A career interview dialogue system using large language model-based dynamic slot generation. In Owen Rambow, Leo Wanner, Marianna Apidianaki, Hend Al-Khalifa, Barbara Di Eugenio, and Steven Schockaert, editors, *Proceedings of the 31st International Conference on Computational Linguistics*. Association for Computational Linguistics, Abu Dhabi, UAE, pages 1562–1584. https://aclanthology.org/2025.coling-main.106/.

Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. Towards emotional support dialog systems. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Online, pages 3469–3483. https://doi.org/10.18653/v1/2021.acl-long.269.

Zhiyang Qi, Takumasa Kaneko, Keiko Takamizo, Mariko Ukiyo, and Michimasa Inaba. 2025. Kokorochat: A japanese psychological counseling dialogue dataset collected via role-playing by trained counselors. https://arxiv.org/abs/2506.01357.

### Biographical sketch

Ekai Hashimoto is a PhD student at the Graduate School of Nagoya Institute of Technology. Prof. Shun Shiramatsu supervises him. His work focuses particularly on the practical application of large language models (LLMs) in dialogue design.