

Towards IoT Workflows for Kitchens Enabled with Ambient Intelligence: A Position Paper

Filippos Ventirozos^{1,2}, Riza Batista-Navarro^{2,*}, Bijan Parsia² and Sarah Clinch²

¹Department of Computing and Mathematics, Manchester Metropolitan University, Oxford Road, Manchester M15 6BH, UK

²Department of Computer Science, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Abstract

The Internet of Things and ambient intelligence are together enabling smart environments that can anticipate and meet user needs. However, Mark Weiser's fully-fledged vision for ambient intelligence is still a long-held dream. This article explores user requirements for creating an ambient intelligent kitchen. Cooking can involve complex interactions between IoT devices and users, making it challenging to design effective IoT workflows. We argue that a successful IoT workflow for cooking should be context-aware, offer varying levels of automation, and provide detailed guidance, among other things. This paper then explores natural language processing technologies, along with an end-user development framework, to fulfil these requirements and make the vision attainable, thereby empowering users to control their smart kitchens.

Keywords

Internet of Things, Ambient Intelligence, Smart Kitchens, End-User Development, Meta-design, Natural Language Processing, Workflows

1. Introduction

The embedding of networked sensors and actuators into our surroundings (i.e., the Internet of Things, or *IoT*), together with associated processing, is enabling novel applications at work, in cities and at home. Collected data can be used to understand the current state and to trigger actions intended to meet context-specific needs, creating "smart" environments, interchangeably referred to as environments with *ambient intelligence* (AmI) in this paper. AmI is a variant of artificial intelligence, in which a "smart environment" anticipates and transparently meets its users' needs [1]. Although useful in a wide range of domains, one common vision for this kind of technology is the realisation of smart homes.

The kitchen has a critical role in household activity, particularly in enabling meal preparation. Although individuals now cook both less frequently and for shorter durations, eating at home remains a significant domestic activity [2, 3]. Home cooking comes with a number of associated benefits including reduced household expense, increased nutrition, reduced carbon footprint and opportunities for to connect with family members.

It is estimated that almost 50% time spent on cooking will be automated by the next 10 years [4]. However, a smart kitchen environment can be highly diverse in contrast with other household ones. Consider, for example, a simple lighting controller. Such a device determines the need for light (checking for motion or occupancy, general light levels, and perhaps relevant activities such as watching TV) and then acts to meet the need. This works well when the user behaviour meshes well with the set of actions and triggers. For pre-configured sets of IoT devices, this entails that the behaviours are sufficiently similar so as to fit the set of triggers appropriately. However, kitchen use is highly diverse [5] and cooking itself often requires complex chains of actions (measuring in different ways, mixing at different speeds at different times, heating in stages), making it challenging to ensure that technology successfully anticipates and meets the needs of its users.

AAPEI '24: The First International Workshop on Adjustable Autonomy and Physical Embodied Intelligence, October 20, 2024, Santiago de Compostela, Spain

*Corresponding author.

✉ f.ventirozos@mmu.ac.uk (F. Ventirozos); riza.batista@manchester.ac.uk (R. Batista-Navarro);

bijan.parsia@manchester.ac.uk (B. Parsia); sarah.clinch@manchester.ac.uk (S. Clinch)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

We hypothesise one of the barriers to ambient kitchen adoption is the ability of the end-users to have understanding, control and agency on their IoT workflow—the backbone of supporting operations in the kitchen for the end-user. We propose that much of the “heavy lifting” can be done with natural language processing (NLP) for constructing IoT workflows. The majority of modern recipes, regardless of their heterogeneous forms, contain instructions for completing a process in a series of steps that involve food ingredients, utensils and devices. For cooking devices, these recipes will also provide information about the specific type of device, temperature and duration (“*Heat oven to 220° C [...] Bake for 25-30 mins*”).

In this position paper, we explore the “ideal” IoT workflow from the perspective of the end-user. We conducted a small set of qualitative interviews to obtain information on how smart devices may support different types of users in their cooking activities. We then propose the application of NLP, as well as computer vision, to retrieve cooking knowledge to form the basis of cooking IoT workflows. Furthermore, we discuss how end-user development (EUD) meta-design principles can augment the IoT workflow.

2. Understanding User Requirements in the Kitchen

In order to gain an understanding of the needs of potential end-users of kitchens enabled with ambient intelligence, we interviewed five participants selected based on convenience sampling [6]. It was ensured that within the set of participants, the three primary personas in the kitchen identified by Kerr et al. [5] are represented, described in Figure 1: the “Beginner”, the “Experienced Challenger” and the “Family Chef”.



Figure 1: Three primary kitchen personas: the beginner, the experienced challenger, and the family chef (adapted from Kerr et al. [5]).

2.1. Participant Profiles

Participants were recruited via email using mailing lists and were supplied with a participant information sheet. Each participant was asked a series of demographic questions, including questions about participants’ cooking habits. Subsequently, participants were presented with a set of six questions that were formulated based on key themes identified in the study by Kerr et al. [5]. These themes centred specifically on the cooking process, exploring areas where smart devices might offer valuable assistance. In addition, participants were asked two questions in relation to ambient intelligence in the kitchen, which were framed as “what if” scenarios [7].

Our participants were all adults aged 30+ who cooked for at least two days a week. All participants reported engaging in some form of novel cooking practices (i.e., being adventurous in cooking) although participants differed in how they realised this, e.g., by tweaking known recipes, following new recipes online, or trying their own creations. Such practices were most common for our one beginner participant, most likely due to the fact that they had a very limited set of established cooking behaviours.

Participant personas were derived by the lead author based on a combination of closed-question responses and responses to the open questions described above. This approach allowed us to differentiate between someone who cooks often and with experience, and someone who cooks often but with limited

confidence and experience. Similarly, concerns about cooking duration were used alongside information about the number of children in the household to identify one participant as a family chef.

Out of our five participants, it was determined that three are experienced challengers, one is a beginner and another one is a family chef. Table 1 presents a summary of the participants' characteristics based on their responses to the questions on demographics and cooking habits.

Table 1

A summary of the participants' responses to our questions on demographics and cooking habits. Their corresponding personas are also indicated.

	Age	Cooking Frequency	Novelty in Cooking	# Adults (# Child.)	Persona
1	40-50	7 days/week	monthly	1	Exp. Challenger
2	40-50	7 days/week	weekly	1	Exp. Challenger
3	50-60	2-3 days/week	at least once a month	2 or more	Exp. Challenger
4	30-40	2 days/week	twice a month	5 (3)	Family Chef
5	30-40	5 days/week	every 2-3 days	1	Beginner

2.2. Thematic Analysis

Based on the participants' responses, we identified various themes that are outlined below.

2.2.1. Following Individual Cooking Steps

Precision of Measurements Several participants raised issues related to the measurement of ingredients. This was true not only for the beginner but also for our three experienced challengers, especially when cooking food using international recipes. For instance, their local supplies and cooking practices prompted them to determine quantities in metric units, but the recipes they tried to follow would use American terms such as "a stick" (for butter), or ambiguous expressions such as "a little bit" or "enough".

All participants suggested a role for their imagined smart kitchen whereby users are provided assistance when it comes to the measurement of ingredients, interpreting ambiguous measurements, conversion between different units, dealing with the lack of measuring equipment, and abstracting over measurements by simply prompting the user to put more (or less) of an ingredient into the dish.

Distractions and Keeping Track Prior research [5] had suggested that family chefs were partially characterised by distractions that arose during cooking. However, our interviews showed that distractions affected all participants. What did distinguish our family chef and beginner (apart from level of experience) was their higher likelihood of facing external distractions (e.g., from family members, chores for the family chef) over cooking-related ones (e.g., finding utensils, preparing ingredients). An experienced challenger also noted that the use of technology to aid their cooking (e.g., using a tablet to display a recipe or to look up information) also led to external distractions (e.g., being notified of new emails).

Our experienced challengers also placed much stronger emphasis on distractions that emerged as part of the cooking process. These often emerged as a result of parallel activities that were either inherent in the cooking process itself, or that came as a result of working through a dish without planning ahead (e.g., beginning to cook some ingredients without having prepared those needed in later stages).

All participants agreed that distractions adversely affected their cooking, causing forgetfulness about steps or the need for repeated information checks. Some employed coping strategies, like tasting for missing ingredients, but these were not always effective.

Participants suggested that automated prompts tailored to current situations could help them refocus, especially against external distractions. They envisioned various state-tracking tools, ranging from progress-monitoring cameras to an "AI spoon" that could taste and adjust food seasoning or consistency.

Step Presentation All participants used recipes in some form (cookbooks, notes, online videos, blogs) but with varying preferences. The beginner and family chef favoured online videos. Nonetheless, everyone found them challenging to follow in real-time, requiring frequent pausing and rewinding.

Some participants preferred text and images but noted layout issues. One experienced challenger suggested that audio dictation combined with images would be valuable.

During the interviews, even without being prompted to discuss technology-based interventions, participants made references to conversational agents as a means for obtaining feedback to handle uncertainty, e.g., to seek advice on particular cooking steps.

General Cooking Knowledge Participants faced challenges in following recipes due to insufficient guidance. Our beginner participant sought more advice on adjusting flavours, such as spice levels, while the family chef had difficulty replicating popular dishes accurately. Even non-beginners frequently turned to the web for extra information or clarifications. The online information accessed included ingredient substitutions, user reviews and technique demonstrations. Lastly, one experienced challenger emphasised the importance of timely information retrieval.

2.2.2. Dependencies between Cooking Steps

The family chef suggested that at times the optimal ordering of steps may not be correctly represented in a recipe, whilst all the experienced challengers identified challenges in determining when time periods had elapsed (particularly when these were expressed as event occurrences, e.g., “when the onions have browned”). Overall, determining how long it takes an end-user to follow a particular step—and how this might affect the succeeding steps—is vital in following a recipe workflow. Participants acknowledged the need for ambient kitchens to know the dependencies between cooking steps.

Finally, a participant highlighted potential safety benefits of an imagined smart kitchen, e.g., for preventing cross-contamination, where an end-user could be told to wash certain utensils used with certain ingredients, before using them in a succeeding step.

2.2.3. Record Keeping and Sharing

All our participants identified some value in keeping records within the kitchen. This includes taking notes when they decide to change an ingredient during cooking, or to record the outcome of a certain deviation; these are helpful when adapting dishes to individuals’ tastes, or to note dishes that emerged from unintended changes to the cooking process. All participants stated that they will find value in having a kitchen that can automatically log their activities both for personal use and for sharing with others.

2.2.4. Personalisation

Our participants also identified a number of other opportunities for personalisation in a smart kitchen, with adaptations to ingredients (to cater to food preferences) and kitchen equipment. In the case of the former, participants envisaged systems that would adjust quantities in accordance with personal preferences (e.g., for seasoning, spices or texture), scaling a recipe to household size, or to make more significant dish adaptations in line with dietary requirements. With respect to kitchen equipment, participants frequently suggested that cooking times could be adjusted to account for differences in oven temperature, either by altering instructions given to the user or by having the kitchen itself control the cooking process.

2.2.5. Agency

Agency refers to giving users control over the degree of automation, allowing fallback to user control in cases of uncertainty. That is, users can choose which parts of the recipe workflow they want to modify, execute or carry out manually.

Our participants had very different expectations about the degree to which the kitchen itself would intervene in cooking tasks. In many cases, participants simply expected that the smart kitchen would make their existing processes more efficient by providing additional information or prompts. However, some participants also saw a role for technology to automate simple manual tasks (e.g., adjusting oven temperature) through to very high levels of automation (i.e., equivalent to having access to a physical assistant). However, as envisaged levels of automation increased, so too did concerns about agency or the potential loss of an enjoyable personal cooking practice. We note that for our beginner participant, neither of these two concerns were raised, but they did express scepticism that (despite their desire for one) a fully automated kitchen might be unrealistic in the short term because of the importance of personal judgement and creativity in the cooking process. Lastly, the beginner participant raised concerns that automation may not respond promptly or appropriately in unexpected circumstances (e.g., if something catches fire), leading to increased risk.

3. Technological Recommendations: AI and EUD

In this section, we provide recommendations on how the end-user requirements outlined in the previous section can be addressed through a combination of AI and IoT workflows.

3.1. Understanding Recipes through Natural Language Processing

Natural language processing (NLP) provides an effective way to extract procedures and measurements from cooking recipes. We examine relevant literature, including multi-modal approaches, to automatically generate IoT workflows solely from these recipes.

End-user needs in relation to *precision of measurements* can be partially met by the NLP task of named entity recognition (NER) [8]. However, recipes often use expressions that are relative rather than absolute (e.g., “fill until water covers the pot”). Also, they do not always specify all the ingredients necessary in each step, since readers are expected to use their common sense knowledge in interpreting the steps. More recent studies have explored the use of multi-modal data to supplement ingredients and tools omitted in text, e.g., by using any images shown for each cooking step [9]. Hence, it is evident that research undertaken in analysing text and images can provide end-users more realistic measurements.

Previously reported NLP studies employed a variety of transformer-based models [10, 11, 12] to infer how the state or location of an entity (ingredient) changes as it undergoes a process. For example, seminal work by Dalvi et al. [12] assembles a state change matrix (Move, Create, Destroy, No change) from a series of instructional steps (including those from recipes), building a graph that denotes which actions depend on which. Advances in this domain, including multi-modal approaches (i.e., adding information from images or videos) could resolve participants’ needs for *keeping track* of the state of various ingredients involved in a recipe.

The computer vision (CV) tasks of action recognition and object detection for the cooking domain [13, 14, 15, 16] could aid in *step presentation* by enabling a conversational agent to give prompts to a user when necessary. Such agents, largely underpinned by large language models (LLMs) [17], have been shown to be capable of interactively communicating with end-users [18]. In addition, combining action recognition with conversational agents has been tested in the cooking domain [19]. Thus, existing approaches within both the NLP and CV literature could accommodate presentation requirements.

General cooking knowledge can be addressed by NLP through the employment of pre-trained language models [18]. These models were trained on a massive amount of documents, including cooking recipes. These type of models could address participants’ needs when seeking further information regarding a cooking step or an ingredient, e.g., through question answering.

A number of NLP studies [20, 21, 22, 23] focussed on extracting a workflow of cooking actions and their corresponding ingredients or cooking tools. To some extent, this could help users in dealing with *dependencies between cooking steps*.

Given the above, it should be feasible to parse a textual recipe into an IoT workflow for automating kitchen tasks, that specify how each device should be configured or operated (e.g., in terms of mode, start

time, duration and temperature). However, despite the most recent advances in NLP and in multi-modal modelling, challenges persist due to the nature of recipe data. Instructional text is usually not precise and some measurements and configurations (including timings, device settings and quantities) still need to be inferred, and may vary depending on each user’s available devices and preferences. For instance, the instruction “*bake for 30 mins, until golden brown*” will require the inference of the following steps on the part of an agent with ambient intelligence:

- setting the temperature
- opening the door
- expecting the dish inside
- performing a check as to whether the dish is golden brown or if 30 minutes have passed, and stopping the process if so
- alerting the user to the next step.

To some extent, these issues are mitigated by constructing recipe databases with precise information. For instance, process-oriented case-based reasoning can suggest recipes to end-users who come with very specific queries in relation to number of calories or type of cuisine [24]. However, such databases require recipes to have been already parsed into a structured format, which poses a hindrance to large-scale automation. Furthermore, in case-based reasoning, the conveyed cooking actions are directed towards the user rather than devices or machines.

Although it is evident that NLP can support the construction of IoT workflows for a kitchen with ambient intelligence, to our knowledge, NLP for recipes remains relatively under-explored. Furthermore, there exist no testbeds for ambient intelligent kitchens that allow for incorporating NLP models for extrinsic evaluation.

3.2. End-User Development

The field of EUD places particular emphasis on *personalisation*. This dimension is not addressed through natural language processing of recipes, as they are generally based on collective insights from community members such as cooks and food bloggers, rather than individual user preferences. In the EUD paradigm, and more broadly in end-user programming, users can specify rules to govern system behaviour in response to particular events or contexts. For example, Ghiani et al. [25] offered mechanisms for users to set natural language rules that dictate the behaviour of kitchen appliances, such as turning off the oven when the user leaves the house.

Our interviews indicated that users envision *agency* over a variety of assistive devices with different levels of autonomy. In contrast, existing research on ambient kitchens usually focusses on a limited set of devices with predefined functionalities [26, 27, 28]. Furthermore, commercial products that we are aware of, such as Thermomix TM6¹, Bosch Cookit², TOKIT³, and Xiaomi Smart Cooking Robot⁴, primarily offer smart devices with a range of pre-configured functionalities. However, they lack multi-device collaboration and proactive personalisation features.

Despite significant progress in the field, we agree with the findings presented by [29]: there are still barriers to achieving collaboration between multiple (kitchen) devices. Additionally, there is a notable absence of adaptive integration for new devices and features, particularly those that can proactively sense and respond to end-user needs.

We propose the construction of IoT workflows that leverages existing recipe mining technologies using NLP, augmented by an EUD meta-design framework. According to Fischer’s definition [30]:

“Meta-design is a conceptual framework for defining and creating social and technical infrastructures in which new forms of collaborative design can occur.”

¹<https://www.thermomix.com/tm6>

²<https://cookit.bosch-home.com/de/faq/>

³<https://uk.tokitglobal.com/>

⁴<https://www.mi.com/de/product/xiaomi-smart-cooking-robot/>

Inspired by Ventirozos et al. [31], we envision a hierarchical meta-design that allows stakeholders with varying degrees of technical expertise to collaborate. The aim is to integrate lower-level semantic representations, such as Petri nets and behaviour trees, with higher-level IoT EUD programming styles, including natural language and block-based visual programming languages. This approach empowers end-users with their preferred tools for crafting their own workflows, thereby reducing the burden on professionals to design workflows for each new device and recipe. Additionally, it places the control in the hands of the end-users, allowing them to test and tailor their ambient intelligence settings transparently, thereby fulfilling the dual requirements of *personalisation* and *agency* (i.e., determining which functionalities they want to be automated).

Moreover, we acknowledge that the ambient kitchen ecosystem involves multiple stakeholders. These range from IoT device vendors and R&D practitioners (developing new models for NLP, object detection, and action recognition) to chefs, food bloggers, nutritionists and food engineers—all contributing their unique perspectives. Our EUD meta-design vision aims to democratise the ambient intelligence space by enabling these various stakeholders to participate in the development of ambient kitchens, even if they possess limited technical knowledge.

To illustrate our vision, consider the following scenario demonstrating the future utility of kitchens enabled with ambient intelligence:

Jo plans a festive dinner for the family. The grandparents want their favourite dishes, so Jo decides to cook them by reusing her recipe workflows from last year. The kids, however, want a new recipe they saw online. Also, they want to cook it with their grandma. The AmI kitchen agent suggests the best order to implement the recipes between Jo and the grandma with the kids. Once the cooking starts, the agent reminds grandma what to put in and observes that the utensils are cleaned properly between cutting the meat and the veggies. The kids ask the AmI agent where mango grows. The agent provides answers and videos to educate and shows all the recipes that they can try with it. Suddenly, the bell rings and the uncle comes and brings an air fryer. Jo is thinking that the new device can alleviate some of the work of the other devices, so she asks the agent to incorporate the air fryer into the recipes. The agent looks in its knowledge base for parts of the recipes that can be done with the air fryer and shows them to Jo. Jo decides that is preferable to not have the grandparents' recipe on the air fryer since they might not be used to it, so she informs the agent about this by voice and also says that she wants to automate most of the recipes since there is a lot to prepare. After the big dinner, Jo sees that the combined IoT workflow is a success, so it is shared with friends. One of the friends is a restaurant owner, who liked one of the recipes with the air fryer; she reuses the workflow and asks her partner to inspect it and edit it using his preferred workflow coding tool for their semi-automated kitchen.

4. Discussion & Limitations

Our investigation of technological recommendations is by no means exhaustive and a larger interviewee sample would potentially elicit further requirements. Nevertheless, our study indicates that domain requirements for an AmI kitchen could be satisfied with NLP and EUD technologies to some extent (with some limitations in terms of cohesion and evaluation). Moreover, even from a small sample of participants, the need for a framework to support agency (and varying degrees of automation) and personalisation was evident. We suggest further research in EUD meta-design to create new inclusive environments for human collaboration and augmented intelligence to democratise and speed up the development of AmI kitchens.

5. Conclusion

In this position paper, we sought to enrich understanding of different kinds of users' cooking behaviours and their associated requirements for a kitchen enabled with ambient intelligence (AmI kitchen). We

demonstrate that NLP and other relevant AI methods can satisfy most of the requirements, and propose the use of EUD meta-design to complement the aforementioned approaches to fulfil the agency and personalisation requirements needed for user adoption.

Acknowledgments

We would like to thank ARM Ltd and the UK EPSRC under grant number EP/S513842/1 (studentship 2109081), whose funding made this study possible.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] A. Vasilakos, W. Pedrycz, *Ambient Intelligence, Wireless Networking, And Ubiquitous Computing*, Artech House, Inc., USA, 2006.
- [2] A. Gatley, M. Caraher, T. Lang, A qualitative, cross cultural examination of attitudes and behaviour in relation to cooking habits in france and britain, *Appetite* 75 (2014) 71 – 81. URL: <http://www.sciencedirect.com/science/article/pii/S0195666313005011>. doi:<https://doi.org/10.1016/j.appet.2013.12.014>.
- [3] G. Ma, Food, eating behavior, and culture in chinese society, *Journal of Ethnic Foods* 2 (2015) 195–199. URL: <https://www.sciencedirect.com/science/article/pii/S2352618115000657>. doi:<https://doi.org/10.1016/j.jef.2015.11.004>.
- [4] V. Lehdonvirta, L. P. Shi, E. Hertog, N. Nagase, Y. Ohta, The future(s) of unpaid work: How susceptible do experts from different backgrounds think the domestic sphere is to automation?, *PLOS ONE* 18 (2023) 1–16. URL: <https://doi.org/10.1371/journal.pone.0281282>. doi:10.1371/journal.pone.0281282.
- [5] S. J. Kerr, O. Tan, J. C. Chua, Cooking personas: Goal-directed design requirements in the kitchen, *International Journal of Human-Computer Studies* 72 (2014) 255–274. URL: <https://www.sciencedirect.com/science/article/pii/S1071581913001365>. doi:<https://doi.org/10.1016/j.ijhcs.2013.10.002>.
- [6] S. J. Stratton, Population research: Convenience sampling strategies, *Prehospital and Disaster Medicine* 36 (2021) 373–374. doi:10.1017/S1049023X21000649.
- [7] C. Potts, K. Takahashi, A. Anton, Inquiry-based requirements analysis, *IEEE Software* 11 (1994) 21–32. doi:10.1109/52.268952.
- [8] S. R. Gunamgari, S. Dandapat, M. Choudhury, Hierarchical recursive tagset for annotating cooking recipes, in: *Proceedings of the 11th International Conference on Natural Language Processing, NLP Association of India, Goa, India, 2014*, pp. 353–361. URL: <https://aclanthology.org/W14-5149>.
- [9] Y. Zhang, Y. Yamakata, K. Tajima, Supplementing omitted named entities in cooking procedural text with attached images, in: *2021 IEEE 4th International Conference on Multimedia Information Processing and Retrieval (MIPR), 2021*, pp. 199–205. doi:10.1109/MIPR51284.2021.00037.
- [10] L. Zhang, H. Xu, A. Kommula, C. Callison-Burch, N. Tandon, OpenPI2.0: An improved dataset for entity tracking in texts, in: Y. Graham, M. Purver (Eds.), *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, St. Julian’s, Malta, 2024, pp. 166–178. URL: <https://aclanthology.org/2024.eacl-long.10>.
- [11] N. Tandon, K. Sakaguchi, B. Dalvi, D. Rajagopal, P. Clark, M. Guerquin, K. Richardson, E. Hovy, A dataset for tracking entities in open domain procedural text, in: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational

- Linguistics, Online, 2020, pp. 6408–6417. URL: <https://aclanthology.org/2020.emnlp-main.520>. doi:10.18653/v1/2020.emnlp-main.520.
- [12] B. Dalvi, N. Tandon, A. Bosselut, W.-t. Yih, P. Clark, Everything happens for a reason: Discovering the purpose of actions in procedural text, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 4496–4505. URL: <https://aclanthology.org/D19-1457>. doi:10.18653/v1/D19-1457.
- [13] S. Bansal, S. Khandelwal, S. Gupta, D. Goyal, Kitchen activity recognition based on scene context, in: 2013 IEEE International Conference on Image Processing, 2013, pp. 3461–3465. doi:10.1109/ICIP.2013.6738714.
- [14] J. Monteiro, R. Granada, R. C. Barros, F. Meneguzzi, Deep neural networks for kitchen activity recognition, in: 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2048–2055. doi:10.1109/IJCNN.2017.7966102.
- [15] M. Ramadan, A. El-Jaroudi, Action detection and classification in kitchen activities videos using graph decoding, The Visual Computer 39 (2023) 799–812. URL: <https://doi.org/10.1007/s00371-021-02346-5>. doi:10.1007/s00371-021-02346-5.
- [16] I. Azurmendi, E. Zulueta, J. M. Lopez-Guede, J. Azkarate, M. González, Cooktop sensing based on a YOLO object detection algorithm, Sensors 23 (2023). URL: <https://www.mdpi.com/1424-8220/23/5/2780>. doi:10.3390/s23052780.
- [17] S. Schöbel, A. Schmitt, D. Benner, M. Saqr, A. Janson, J. M. Leimeister, Charting the evolution and future of conversational agents: A research agenda along five waves and new frontiers, Information Systems Frontiers 26 (2024) 729–754. URL: <https://doi.org/10.1007/s10796-023-10375-9>. doi:10.1007/s10796-023-10375-9.
- [18] J. I. Choi, S. Kuzi, N. Vedula, J. Zhao, G. Castellucci, M. Collins, S. Malmasi, O. Rokhlenko, E. Agichtein, Wizard of tasks: A novel conversational dataset for solving real-world tasks in conversational settings, in: Proceedings of the 29th International Conference on Computational Linguistics, International Committee on Computational Linguistics, Gyeongju, Republic of Korea, 2022, pp. 3514–3529. URL: <https://aclanthology.org/2022.coling-1.310>.
- [19] R. Kojima, O. Sugiyama, K. Nakadai, Audio-visual scene understanding utilizing text information for a cooking support robot, in: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2015, pp. 4210–4215. doi:10.1109/IROS.2015.7353973.
- [20] J. Jermurawong, N. Habash, Predicting the structure of cooking recipes, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Lisbon, Portugal, 2015, pp. 781–786. URL: <https://aclanthology.org/D15-1090>. doi:10.18653/v1/D15-1090.
- [21] O. Abend, S. B. Cohen, M. Steedman, Lexical event ordering with an edge-factored model, in: Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Denver, Colorado, 2015, pp. 1161–1171. URL: <https://aclanthology.org/N15-1122>. doi:10.3115/v1/N15-1122.
- [22] Y. Yamakata, S. Mori, J. Carroll, English recipe flow graph corpus, in: Proceedings of the Twelfth Language Resources and Evaluation Conference, European Language Resources Association, Marseille, France, 2020, pp. 5187–5194. URL: <https://aclanthology.org/2020.lrec-1.638>.
- [23] D. P. Papadopoulos, E. Mora, N. Chepurko, K. W. Huang, F. Ofli, A. Torralba, Learning program representations for food images and cooking recipes, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 16538–16548. doi:10.1109/CVPR52688.2022.01606.
- [24] R. Bergmann, L. Grumbach, L. Malburg, C. Zeyen, Procake: A process-oriented case-based reasoning framework., in: ICCBR Workshops, volume 2567, 2019, pp. 156–161.
- [25] G. Ghiani, M. Manca, F. Paternò, C. Santoro, Personalization of context-dependent applications through trigger-action rules, ACM Trans. Comput.-Hum. Interact. 24 (2017). URL: <https://doi.org/10.1145/3057861>. doi:10.1145/3057861.

- [26] A. Neumann, C. Elbrechter, N. Pfeiffer-Leßmann, R. Kõiva, B. Carlmeyer, S. Rütter, M. Schade, A. Ückermann, S. Wachsmuth, H. J. Ritter, “kognichef”: A cognitive cooking assistant, *KI - Künstliche Intelligenz* 31 (2017) 273–281. URL: <https://doi.org/10.1007/s13218-017-0488-6>. doi:10.1007/s13218-017-0488-6.
- [27] B. Bouchard, K. Bouchard, A. Bouzouane, A smart cooking device for assisting cognitively impaired users, *Journal of Reliable Intelligent Environments* 6 (2020) 107–125. URL: <https://doi.org/10.1007/s40860-020-00104-3>. doi:10.1007/s40860-020-00104-3.
- [28] G. Kondylakis, G. Galanakis, N. Partarakis, X. Zabulis, Semantically annotated cooking procedures for an intelligent kitchen environment, *Electronics* 11 (2022). URL: <https://www.mdpi.com/2079-9292/11/19/3148>. doi:10.3390/electronics11193148.
- [29] D. Marikyan, S. Papagiannidis, E. Alamanos, A systematic review of the smart home literature: A user perspective, *Technological Forecasting and Social Change* 138 (2019) 139–154. URL: <https://www.sciencedirect.com/science/article/pii/S0040162517315676>. doi:<https://doi.org/10.1016/j.techfore.2018.08.015>.
- [30] G. Fischer, E. Giaccardi, *Meta-design: A Framework for the Future of End-User Development*, Springer Netherlands, Dordrecht, 2006, pp. 427–457. URL: https://doi.org/10.1007/1-4020-5386-X_19. doi:10.1007/1-4020-5386-X_19.
- [31] F. Ventirozos, R. Batista-Navarro, S. Clinch, D. Arellanes, Iot cooking workflows for end-users: A comparison between behaviour trees and the dx-man model, in: *2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, 2021, pp. 341–350. doi:10.1109/MODELS-C53483.2021.00057.

A. Interview Questions

I. Demographics and cooking habits:

1. What age range describes you?
2. How often do you cook?
3. How often do you cook something new?
4. Do you usually cook for more than 5 people?
5. Do you have children, if yes how many?

II. Questions drawn from the study by Kerr et al. [5]:

1. How easy can you follow a recipe? What kind of modality do you prefer and why?
2. Do you anticipate any advances in device configurations? For instance, taking out the guesswork and automatically be turned on/off when necessary.
3. How easy is for you to get advice on specific steps of a recipe, what do you find most cumbersome?
4. If you weigh ingredients, do you find the task irksome? Would automation in weighting be helpful? For instance, the scale knowing what ingredient is being measured and how much is the correct weight - the same applies with liquids.
5. Whenever you have inspirations whilst cooking, do you log them and how? Do you think this process could be improved by an intelligent kitchen knowing what ingredients and device configurations you used?
6. If you share recipes, would the above process also be helpful?
7. Do you experience distractions whilst cooking? What is the most annoying thing about it?

III. Automation and Instruction:

1. Can you imagine yourself cooking by giving instructions to your kitchen, to do the labour intensive tasks? How would you envision that? Would you be worried that the automation could screw some things up?
2. What if Ambient Intelligence could adapt to your needs and your environment? Tailor cooking recipes and steps for you?