Exploring Access to Cooking Information for People with Vision Impairments

Ashley Wang

CMU-CS-23-122

July 2023

Thesis Committee:
Patrick Carrington (Chair)
Jeffrey Bigham

Computer Science Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Submitted in partial fulfillment of the requirements for the degree of Master of Science.

This project is a collaboration between CMU's AXLE Lab and Google's People + AI Research Group. Funding for this project is provided by Google Research.

Keywords: Cooking, Recipes, Blind, Accessibility, AI Assistive Technology

Abstract

Cooking is an essential task that people frequently perform in their daily lives to prepare foods for sustenance, and is a major contributor to improved quality of life. Prior to the cooking process itself, obtaining and preparing ingredients, navigating a kitchen, learning cooking terms, and understanding how to use tools are all necessary skills for a home cook to learn. Recipes play an essential role in this learning process, as they provide guidance for each step of preparing a dish. However, most recipes, especially in digital formats, contain an abundance of visual information without non-visual alternatives. This reliance on visual communication leads to barriers for blind cooks. Therefore, it is important to explore the information need of blind cooks, their interactions with information while cooking, and how recipes and technology should be designed to not solely rely on visual delivery of information.

Toward this goal, we conducted semi-structured interviews with 20 visually-impaired home cooks and 4 cooking instructors at Saavi Services for the Blind, a blindness rehabilitation training center providing life skills training using non-visual techniques. We found that the cooking process of visually-impaired home cooks consists of 5 stages: information searching and comparison, information extraction and manipulation, learning and ideation through information, information interaction, and information identification. In this work, we explore how blind cooks interact with information throughout each stage of cooking that we identified. Furthermore, we uncover unique findings of existing methods, strategies, and challenges in accessing information while cooking (e.g., tracking specific steps while cooking) and present the information need for cooking activities (e.g., tolerance on timing for cooking steps). We also contributed 6 design features for technology aiming to deliver information to visually-impaired people while cooking (e.g., restricted gestures for interaction). Overall, our findings provide a roadmap of information access for people with vision impairments throughout different cooking stages.

Acknowledgments

I would like to express my deepest gratitude to Professor Patrick Carrington, my advisor and committee chair who guided me through this project.

I'm very grateful to Shaun Kane for his invaluable guidance and feedback.

I am extremely thankful to Franklin Mingzhe Li for his help and guidance, especially throughout the user studies and affinity diagramming process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cooking is a fundamental life skill that enables people to prepare meals for both sustenance and enjoyment. However, visually-impaired individuals can find it challenging to cook, and even develop an aversion to it due to difficulties in accessing information (e.g. ingredients, steps, substitutions) and interacting with information (e.g. navigating a recipe) [17, 34, 38]. These barriers cause many people in the visually-impaired community to rely on frozen and takeout foods, which have been shown to be contributors to malnourishment and poor quality of life [30]. To better support blind people in cooking, it is necessary to investigate how visually-impaired individuals access information while cooking, what information they want and do not want, and how recipes and assistive technologies can be designed to deliver information in an accessible way. Our goal is to answer these questions so that visually-impaired home cooks can prepare their meals both independently and confidently, which are important for maintaining a healthy and balanced lifestyle [38].

Toward this goal of exploring information access of visually-impaired home cooks, we chose to use recipes as a probe. Recipes are an integral part of the cooking process, as they not only allow people to receive cooking knowledge for learning purposes, but also allow people to pass on cooking knowledge to others [21, 41]. Importantly, recipes provide guidance for each step of preparing a dish, and are a commonly used learning tool for home cooks of all skill levels. However, many digital recipes are heavily reliant on visual communication (e.g. explaining steps using images) and contain information that is inaccessible or unhelpful to people with vision impairments [38]. For example, many recipes describe the status of food with solely visual descriptors (e.g. cook until golden brown), which assumes its user is sighted. The lack of non-visual alternatives and accessible information in recipes leads to confusion, requiring additional effort for people with vision impairments to follow [17, 34, 38].

In addition to the availability of information, receiving accessible information through a preferred modality while cooking is essential for people with vision impairments [38], especially because many people choose to follow a recipe when cooking in order to ensure they can achieve a desired result. People with vision impairments can leverage multiple approaches, such as screen readers [7, 16] or smart devices, to access recipes from online sources (e.g., Allrecipes [2], Food Network [9], YouTube videos). While some previous research has examined practices of visually-impaired home cooks (e.g., [38]), it is still necessary to investigate what challenges exist when using each information access approach, and what modalities are preferred when accessing information during cooking.

In order to improve the accessibility of cooking for people with vision impairments, it is necessary to first understand what type of information is important to blind people while cooking [38]. Furthermore, as AI tools continue to advance, it is also important to understand how such technology should be developed to deliver this information to blind people while they are cooking. Therefore, we are interested in exploring the following research questions:

- RQ1: What methods and strategies do people with vision impairments employ when accessing information while cooking? What challenges do they encounter?
- RQ2: What types of information are important to blind individuals during the cooking process? Why?
- RQ3: How should technologies be developed to deliver accessible information to blind individuals while cooking?

To explore our research questions, we first conducted a semi-structured interview study with 20 people with vision impairments who have experience cooking for themselves and four cooking instructors from a vision rehabilitation center. We first describe current methods, strategies, and challenges of accessing visual information while cooking (e.g., tracking specific steps while cooking) (Section 4.1). We then describe the information needs of blind people with respect to information content and structure (e.g., tolerance on timing for cooking steps) (Section 4.2). Next, we highlight design features for technologies to deliver accessible information while cooking (e.g., restricted gestures for interaction) (Section 4.3). Finally, we discuss design guidelines and future directions to improve information access for people with vision impairments (e.g., LLMs for Kitchen-AI systems, non-visual database for Blind people) (Section 5).

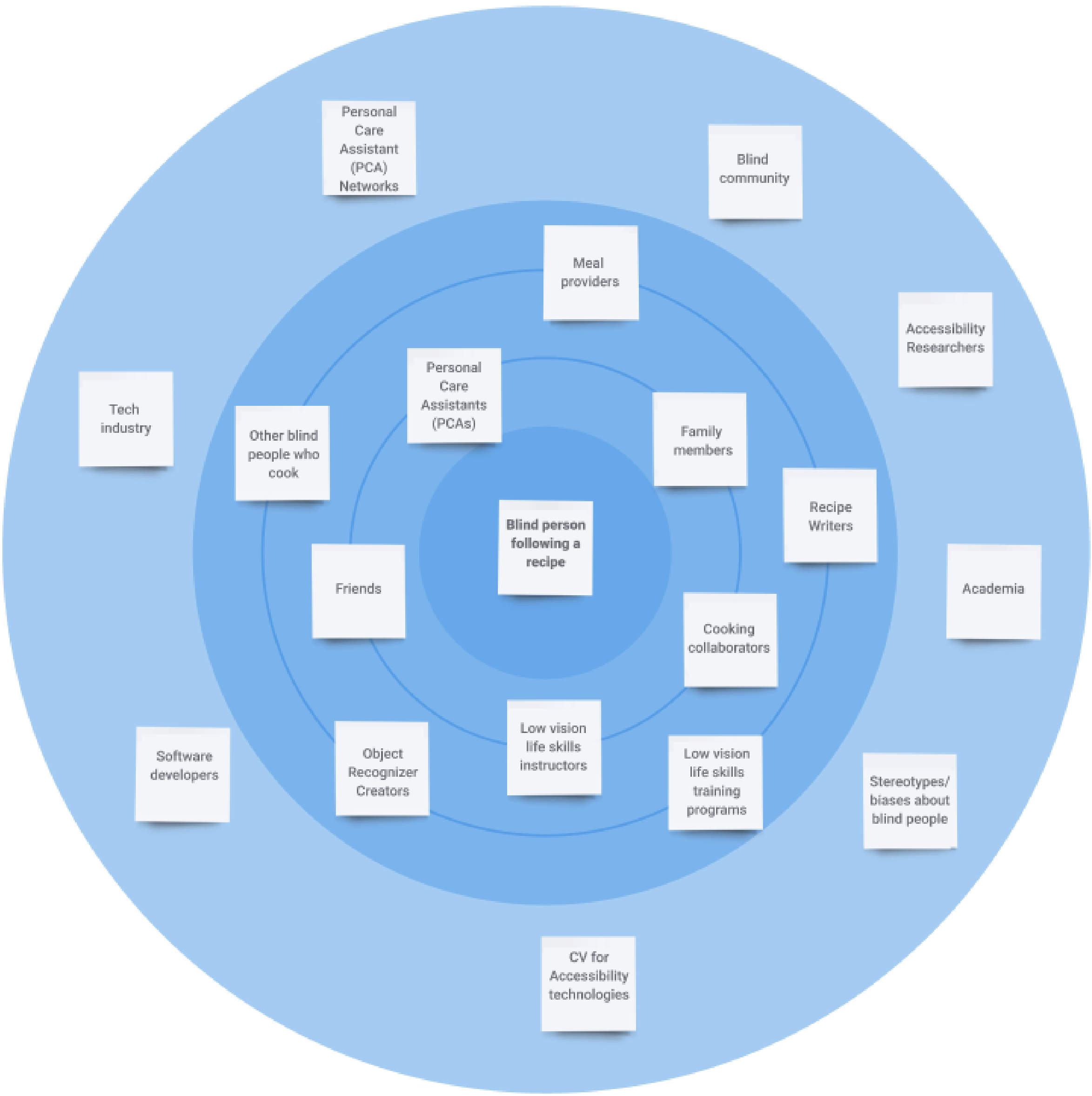

Figure 1.1: Stakeholder map identifying our central stakeholder, people and groups directly impacting the central stakeholder (inner ring), and systems and groups indirectly impacting the central stakeholder (outer ring).

Chapter 2

Related Work

In this section, we first present the importance of cooking to people with vision impairments and their associated experiences. We then introduce different approaches to accessing information through technologies for people with vision impairments. Finally, we discuss prior research on both cooking-related technology and recipe-related technology to reflect how our work bridges the gap of existing technology, and how technology should be developed for people with vision impairments to access information during different cooking stages.

2.1 Cooking Experiences of People with Vision Impairments

Vision impairments have been shown to impact the experiences of food, eating, and cooking [17]. Through surveying over 100 visually impaired people, Jones et al. [30] found a significant correlation between the severity of vision impairment and the difficulty to shop for ingredients and to cook meals. In fact, researchers even found that vision impairments are related to malnourishment and poor quality of life [30]. By conducting semi-structured interviews with nine visually impaired people in 2009, Bilyk et al. [17] found that people with vision impairments heavily rely on prepared food from outside sources. For example, all participants in the study reported consuming at least 40% of their dinners in restaurants, mostly in order to avoid cooking. More recently in 2017, Kostyra et al. [34] found that out of 250 survey respondents, 49.6% prepare their meals independently, while the remaining participants prepare their meals with the help of either sighted and/or blind individuals. When asked about their difficulties with food preparation, 82.1% and 72% found peeling vegetables and frying foods to be difficult, respectively. In contrast, participants reported that activities not involving heat or tools were easier, such as: preparing sandwiches and washing fruits. Because of the ease of preparation, 57.6% of participants bought ready-to-eat products, although only 14% bought ready-to-heat meals [34].

Prior research uncovered various ways to support people with vision impairments to access information non-visually while cooking. In a previous study by Yoshikawa et al. [52], crispiness was determined to be one of the key sensations produced by foods when eating them. Zampini et al. [53] found that the perception of the crispiness and staleness of potato chips can be altered by modifying the auditory cues given to participants when eating them. Li et al. [38] also found multisensory contributions to food perception by observing cooking recordings of visually impaired people; they found that visually impaired cooks heavily rely on their auditory and tactile senses. For example, they often like to spread jam or butter with their fingers to make sure it is evenly distributed [38]. They also use sound to determine if water is done boiling or not [38]. It is still unknown what information is important to visually impaired people when they are cooking, and how they would like this information delivered to them.

2.2 Information Access by People with Vision Impairments

To access information in daily activities, people with vision impairments use screen readers [28, 37, 40, 44], speech synthesizers [18, 47], and Braille displays [13, 45] to theoretically have equal online access to information as sighted people. However, they are limited by the information that is available, and not what they are able to read [13]. The issue of making information accessible to the visually impaired community is an issue of equity and standards on a larger scale [13]. With regards to cooking information, Bilyk et al. [17] found in 2009 that the primary sources of food-preparation instruction that visually impaired individuals use are: their mothers, resources from a non-profit organization, and high school cooking classes. In addition, they identified 11 obstacles related to cooking, including: difficulty accessing recipes, measuring ingredients, and frying/browning food. It is unknown what the best way of addressing these challenges is.

Now, with the involvement of technology in all aspects of life, there are more ways of accessing information for cooking, including various online resources to find recipes (e.g., AllRecipes.com [2], FoodNetwork.com [9]), as well as various ways of accessing food-preparation instructions (e.g., Alexa [3], YouTube [1]). There are also recipes and personal blogs on which users can post any recipes that they wish to share [5]. However, it is unknown how people with vision impairments leverage these different approaches and the information need from them while cooking.

2.3 Cooking-related Technology

With advances in sensors, wearable technology, and smart appliances, it is important to explore what technology exists to help people cook in the kitchen. Prior research has been done to explore the procedures and techniques involved in cooking [32, 35, 46]. In an effort to understand and categorize what happens to food as it cooks, Kato and Hasegawa proposed an interactive cooking simulator system that provides users with feedback regarding the physical and chemical food changes occurring during cooking [32]. This system helps people understand cooking operation effects, and could also be extended to provide detailed non-visual explanations for the visually impaired. Furthermore, Arakawa et al. [14] developed PrISM-Tracker, which is a multimodal framework using wearable sensors that improves upon existing human activity recognition models in order to guide people in performing procedural tasks such as cooking or setting up an appliance.

In terms of cooking technology for people with vision impairments, prior research explored using computer vision [23, 24, 42, 49], voice interactions [12, 20], and tactile marking [25, 27] to better support people with vision impairments interacting with different kitchen appliances.

For example, Guo et al. [24] leveraged computer vision and crowd workers to support people with vision impairments to interact with different kitchen appliances. Beyond making appliance interfaces accessible, prior research has also explored various approaches support people with vision impairments to interact with devices through different gestural interactions (e.g., [15, 31, 39]). However, it is unknown how these technologies can be adopted and designed to better support people with vision impairments accessing information under different cooking stages.

2.4 Recipe-related technology

Recipes act as the main source of cooking activities [19]. Prior research has been conducted in order to explore recipe wording and construction. For example, Tasse and Smith [48] created CURD, a database of recipes annotated using the MILK language, a meaning representation language based on first-order logic, to break recipes and steps into pieces of a specific format and function. Various online services that can recommend a recipe based on a set of ingredients exist, such as the Kroger Chefbot [8] and ChatGPT [6], although research into the quality of the provided recipes have not been performed.

To provide more adaptive recipes based on individual needs, Teng et al. [50] measured the role of ingredients in recipes, as well as the relationships between co-occurring ingredients. They constructed networks to capture these relationships and used these networks, as well as user feedback, to create recipe recommendation algorithms. Instead of looking at ingredients as previous studies did, Kusu et al. [35] calculate the difficulty of recipes based on the cooking activities involved. After extracting cooking activities from recipes and categorizing them into four difficulty levels, they proposed formulas that related cooking activity difficulty level with recipe skill level [35]. Their results show that their approach makes it easier for participants to find suitable recipes based on their skill level. Moreover, Chang et al. [21] created RecipeScape, an interactive system that, given a dish, can search through hundreds of recipes for that dish and categorize them according to cooking approach. Given these approaches on extracting recipe information and provide adaptive and customized recipe, it is unknown how these can be leveraged to people with vision impairments, and more importantly, what are the information needs to people with vision impairments and why.

Chapter 3

Methodology

First, we conducted semi-structured interviews (SSIs) with visually-impaired people possessing some level of cooking experience. The purpose of these SSIs was to understand their existing methods, strategies, and challenges when accessing information (e.g., a recipe) while cooking; information need for people with vision impairments while cooking; and design features for technologies to better deliver accessible information while cooking. To approach our research questions from a different perspective, we conducted SSIs with cooking instructors for the blind from a vision rehabilitation center. Through these interviews, we gained further insight into the challenges that blind people experience when learning too cook, the important information necessary to support blind people when cooking, and how technology can help better support blind learning experiences regarding cooking. Participants were compensated with a $20 Amazon gift card. The recruitment and study procedure was approved by the Institutional Review Board (IRB).

3.1 Interview with Blind People with Cooking Experiences

We recruited 20 blind participants (P1 - P20) (Table 3.1) through various online platforms (e.g., Twitter, Reddit) and email lists (e.g., National Federation of the Blind). The prerequisites to participate in our study were: participants needed to be at least 18 years of age, to be legally or totally blind, have experience with cooking, and be able to communicate in English. Among the 20 participants we recruited, 11 of them were female, and 9 were male (Table 3.1). They had an average age of 43 (SD = 14.1). Nine of them are legally blind, and 11 are totally blind. Each interview took around 75 to 100 minutes to complete.

3.1.1 Pre-study Survey

Before the actual interview, we asked our participants to complete a pre-study survey for their demographic information (vision condition) to ensure they met the prerequisites for the study. In this survey, we also asked about what sources they use to find recipes, as well as three dishes they would like to learn to cook. We used this information as a probe to understand what type of information is important to blind individuals while cooking.

Participant	Age	Gender	Vision Description/Level	Years Cooking	Cooking Frequency
P1	32	Female	Totally Blind	15 Years	3 Times a Week
P2	61	Male	Totally Blind	53 Years	4,5 Times a Week
P3	48	Female	Totally Blind	20 Years	Everyday
P4	57	Female	Totally Blind	40 Years	Twice a Week
P5	47	Male	Legally Blind	23 Years	Several Times a Week
P6	38	Male	Completely Blind Right, Legally Blind Left (Tunnel Vision)	32 Years	Several Times a Week
P7	44	Female	Totally Blind with Light Perception	20 Years	4,5 Times a Week
P8	26	Male	Legally Blind	20 Years	2,3 Times a Week
P9	31	Female	Legally Blind	18 Years	3 Times a Week
P10	61	Female	Totally Blind	45 Years	Everyday
P11	38	Male	Totally Blind	30 Years	Everyday
P12	67	Male	Legally Blind	49 Years	Everyday
P13	36	Male	Legally Blind	Since Teenager	Everyday
P14	30	Female	Totally Blind	11 Years	Once a Month
P15	70	Female	Totally Blind	58 Years	Everyday
P16	24	Female	Legally Blind	2 Years	3,4 Times a Week
P17	39	Female	Totally Blind	26 Years	Everyday
P18	48	Female	Some Light Perception	40 Years	3,4 Times a Week
P19	46	Male	Totally Blind	29 Years	Everyday
P20	24	Male	Legally Blind	14 Years	3 Times a Week

Table 3.1: Participants Demographic Information

3.1.2 Demographic Background [$\sim$ 5 Minutes]:

In our SSIs, we first asked about the demographic background of our participants, which included: age, gender, and vision level.

3.1.3 Methods, Strategies, and Challenges of Accessing Information While Cooking [$\sim$ 20 Minutes]:

We asked our participants about their common practices in accessing cooking information. For example, we asked about the technologies and devices they use to find recipes (digital or physical), how they identify information while cooking, and any interactions they adopt to navigate information while cooking. In addition to our prepared questions, we asked follow-up questions about any interesting points our participants brought up.

3.1.4 Information Need While Cooking [$\sim$ 30 Minutes]:

In the pre-study survey, each participant provided three dishes they would like to learn to cook, as well as their preferred recipe source(s). From their preferred source(s), we selected three recipes for each participant, corresponding to the three dishes for which they expressed a desire to learn. We used this information as a probe to explore each participant's perceptions about information need. For each of the three recipes, researchers read the recipe to the participant, stopping after each step to ask what the participant would like more/less information on. For example, this information could be clarification/specific details on an action, missing instructions, or information that is extraneous or unnecessary.

Figure 3.1: Based on the three dishes and recipe sources provided in the pre-survey study, we selected three recipes for each participant and made a table for each of these recipes in our notes. These tables was used to record any information that our participants wished to add or remove from each recipe step, as well as their general comments about the steps.

3.1.5 Design Features of Technologies Used to Access Information [$\sim$ 30 Minutes]

We also asked participants to describe and discuss their experiences with leveraging technology to access information while cooking (e.g., screen reader, YouTube recipes, AIRA/BeMyEyes). We used these scenarios as probes to further understand the necessary design features of technologies for people with vision impairments in cooking.

3.2 Interview with Blind Cooking Course Instructors

To explore our research questions from a different perspective, we recruited four participants (I1 - I4) who are blind cooking course instructors at Saavi Services for the Blind, a blindness rehabilitation training center providing life skills training using non-visual techniques. (Table 3.2). Among the four instructors we recruited, 3 are female, and 1 is male (Table 3.2). They had an average age of 51 (SD = 18.5). Two of them are legally blind (I1, I3), one is totally blind (I4), and one is sighted (I2). I1, I2, I3, and I4 have 9 months, 10 months, 4 months, and 30 years of experience in teaching blind cooking courses, respectively. Each interview took 60 to 75 minutes to complete.

Participant	Age	Gender	Vision Description	Instruction Duration
I1	58	Female	Legally Blind	9 Months
I2	24	Female	Sighted	10 Months
I3	58	Male	Legally Blind	4 Months
I4	65	Female	Totally Blind	30 Years

Table 3.2: Instructor Demographic Information

3.2.1 Demographic Background [$\sim$ 5 Minutes]:

In the interview with cooking instructors, we first asked about the demographic background of our participants, which included age, gender, descriptions of vision level, and how long they have been instructing blind people to cook.

3.2.2 Experiences and Behaviors of Blind Cooking Instructors [$\sim$ 25 Minutes]:

We asked the instructors about their experiences with instructing blind people in the cooking class. They discussed current teaching practices, existing challenges students encounter in cooking classes, learning barriers, information needs, and communication methods and challenges.

3.2.3 Information Need While Cooking [$\sim$ 15 Minutes]:

We asked the instructors about their perceptions of the information source, information quality, and information delivery for blind people. We also asked them about how information should be designed for blind people, catering to different learning stages and levels of experience.

3.2.4 Design Features of Technologies to Access Information [$\sim$ 20 Minutes]

To conclude the interview, we asked the instructors about the technologies that they use to facilitate learning for their students. Following up on their thoughts about the technologies they use, we asked them for suggestions on technology design for blind people.

3.3 Data Analysis

The semi-structured interviews were conducted through Zoom [29], and all interviews were audio-recorded and transcribed, verbatim. Then, two researchers independently performed open coding [22] on a set of the four same transcripts. They reconciled the codes to form the initial codebook. The two researchers independently coded the remaining transcripts, applying existing codes when appropriate and adding new codes when necessary. They refined the codebook by discussing the codes and resolving conflicts (e.g., missing codes, disagreements). After the two researchers reached a consensus on the final codebook, they performed affinity diagramming [26] using a Miro board [33] to cluster the codes and identify emergent themes.

Chapter 4

Findings

4.1 Methods and Challenges of Accessing Information During Various Cooking Stages

We identified 5 cooking stages based on how a cook's interactions with information change while cooking: information searching and comparison, information extraction and manipulation, learning and ideation through information, information interaction, and information identification. We present the current methods, strategies, and challenges that people with vision impairments have during each of these cooking stages.

4.1.1 Information Searching and Comparison

In our interview with participants, we asked them questions regarding how they search for cooking information (e.g. recipes) and how they make decisions on what to cook. Every participant had a practice of leveraging online resources to learn about non-visual cooking skills (e.g., how to accurately measure ingredients) and about new dishes that are unfamiliar to them. Their commonly-used sources include: recipe websites (e.g., AllRecipes.com, FoodNetwork.com) (20), specialized websites (e.g., VisionAware.org, acb.org) (8), Facebook groups (e.g., Incredible Recipes, The Country Cook) (6), YouTube videos (12), podcasts (5), and smart speakers (e.g., Alexa) (6). P17 describes his experience of leveraging online resources for cooking information:

> "The ads are really frustrating, but once you can get through those, I've had a lot of success. (...) I get the weight of my Turkey and I go search and find all these recipes and things like that, and they've been super helpful and I've even gotten some new ideas that I didn't necessarily know about from reading these recipes and things like that. Get a little more adventurous in my cooking. I'm a great basic cook, but trying new ideas and stuff, recipes are great for that."

A common complaint that our participants had was that online sources contain an abundance of distracting, inaccessible visual content (e.g., ads, commercial videos) that makes navigation difficult. P6 describes his experience with inaccessible online content:

> "When navigating the recipe websites, swipe right to scroll through them. Sometimes that would cause problems. If there was a video in the recipe, it would au-

Stage	Name of Stage	Key Findings
1	Information Searching and Comparison	- Blind and visually-impaired (BVI) people compare information from various sources before choosing a recipe that fits their needs. - Online sources have distracting, inaccessible, visual content (e.g. ads) that make navigation difficult.
2	Information Extraction and Manipulation	- Saving recipes in a text file for future use allows BVI people to customize them and avoid online ads when referring to it later. - Another common way to save recipes for convenient future access is to transcribe video recipes into a text file.
3	Learning and Ideation through Information	- Youtube videos/podcasts are used to learn about new dishes and for inspiration, not to cook along with.
4	Information Interaction	- BVI people heavily rely on tactile senses in cooking, so their hands are constantly covered in food. - Protective coverings (e.g. saran wrap) are placed around devices to prevent food from getting on them. - Smart speakers (e.g. Amazon Alexa) are used for hands-free interactions. - Issues when using smart speakers include: not being able to find exact information, to navigate, and to maintain reading status.
5	Information Identification	- It is especially difficult for BVI people to differentiate between objects that feel the same and to measure liquids. - In order to complete these hard tasks, and to check on the visual characteristics of food, BVI people use crowd- or AI-based vision apps. - Issues with vision apps include: difficulty locating text on a food package, looped readings, and live reading sensitive to camera movement. - Vision apps are unable to reliably read: text over a background picture, tables, and artistic text.

Table 4.1: Key methods used by blind and visually-impaired people when accessing cooking information, and challenges they face during the process. Our findings are organized into 5 cooking stages, which we identified in our study.

tomatically start playing with sound. And even if I were to scroll past it, the video was already playing. Then the sound would just continue. It would continue playing that media. I don't know whether or not the video was still visually on the screen or not. But the audio still. So that was a real endurance, especially when you're you're relying on the audio of talk back at the same time, so I can't remember if it was playing over talk back, or if talk back wasn't working, but I do remember that it basically stopped me in my tracks."

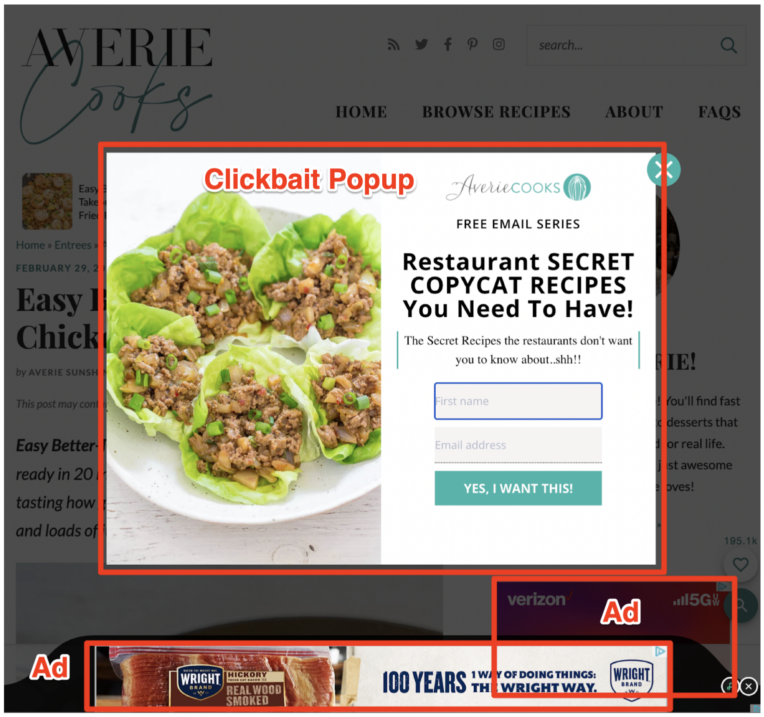

Figure 4.1: An example of a recipe webpage with many ads. Surge AI [11] found that each recipe page within the top search results of Google contains an average of 8.7 ads.

When making a decision, many participants expressed that they compare information obtained from various sources or multiple options from one source. They compared different recipes for the same dish and chose one based on its ingredients, time/effort required, ratings, clarity, and accessibility. P7 explained:

"I honestly just hop on Google, and I start out that way. If you get on Google, the way that it lays everything out, the first thing it will do is show you recipes, and then it shows you a bunch of websites, and then it'll say, "Search results," and it'll show you those recipes again. I'll look to see how many people have rated it. I'll look through three or four recipes about the same thing, and figure out what sounds good to me, what flavors sound right, what timing sounds right. I don't like when people want to overcook something. If it's 10 minutes longer than something else for a steak, I'm going to go the conservative, lower time route. I look through, and figure it out from there, piece together my own idea of what feels like what I want to do."

4.1.2 Information Extraction and Manipulation

To save and keep track of online content for later, we found that our participants prefer saving them either in a Word document or text file. These two options are simple to use, allow for customized editing or note-taking, and allow our participants to avoid online ads and clutter. P1 explains how she saved recipes found in her Facebook groups:

> "So I have joined a bunch of Facebook groups and I'll typically get recipes from those groups, and if I see something I will save it and typically I'll not look at it at the moment. I'll just sort of save and curate and then when I have time, I'll go look at the websites and copy everything out into an easier to read plain text file so that I don't have to deal with ads and web formatting when I'm trying to cook. I need it in a very plain, simple format."

In addition to directly copying-and-pasting the recipe text from websites, our participants also use the approach of manually transcribing and typing information from video recipes onto a text document. P16 reflects on the experience of doing so:

> "I'll just play it [the video] a group of seconds at a time. If she's listing the ingredients, I'll start it and pause it and then type it into my computer and start it and pause it and type it into my computer. It's not convenient, but it's nicer than if I miss an ingredient. I have to watch the whole video again."

4.1.3 Learning and Ideation through Information

From our interviews, we discovered that our participants have a brainstorming/planning stage before deciding what to cook and actually cooking a dish. They expressed that they leverage YouTube and podcasts to ideate dishes they can make. In addition to learning about new dishes, our participants also gained inspiration and enjoyment by witnessing the process of cooking through the audio descriptions, as well as the sound emitted by food, actions, and tools during the cooking process. P20 describes how he benefits from video recipes:

> "In terms of frying, they could just tell you, okay, it's going to sound this way. So you just appreciate the sound of the dish cooking. And then you could also hear whoever is guiding you, what he's saying."

From our interviews with the four instructors (I1 - I4) of blind cooking classes, we learned that people with vision impairments constantly ask questions for confirmation and learning purposes. The instructors also emphasized the importance of engaging in verbal Q&A, rather than hand-over-hand instruction [51] when teaching students, in order to foster independence and active learning. I1 explains his teaching philosophy:

> "We're a structured discovery center, so most of the time it's students having to work through these things on their own. If they have any questions, they can ask me, but usually their questions are always met with the question to help just guide them through the process and think through it themselves. So typically, I avoid... I'm not going to say avoid actually. Typically, I don't do a lot of hand over hand teaching. If I'm giving instruction or if students need any clarification or they're learning something new, I always try to give verbal instruction first."

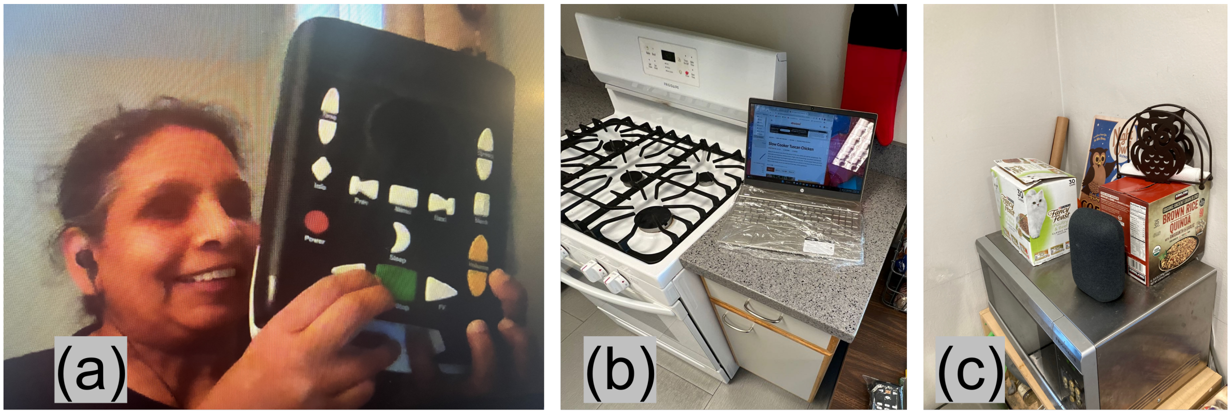

Figure 4.2: Technology used by participants to access recipes. (a) P4 uses a tactile-based audio player to play a recipe. (b) P2 wraps a laptop in saran wrap to prevent it from getting dirty when repeatedly using it with messy hands in the cooking process. (c) P11 uses a smart speaker in the kitchen to get recipe information hands-free.

4.1.4 Information Interaction

We discovered, through our interviews, that participants utilize various methods to access information (e.g. a recipe) while cooking. These methods include: voice-over on their smartphones, screen readers on their laptops, audio players to play text or audio files, smart speakers (e.g. Amazon Alexa, Google Nest), physical cookbooks, video recipes (e.g., YouTube), and smartwatches. Figure 4.2 (a), (b), (c) shows how our participants use these methods in their daily lives.

One common issue that our participants expressed they face while cooking is that their hands are frequently covered in food, as they rely heavily on touch to measure ingredients [38], check for doneness, etc.. However, while following a recipe, our participants need to use and interact with their devices frequently to refer back to the recipe. In order to keep their devices safe from the food on their hands, they either positioned their devices far away from the cooking area and added protective coverings such as plastic wrap (see Figure 4.2 (b)), or opted for hands-free interactions, such as using a smart speaker (see Figure 4.2 (a)). P2 explains how he does this:

> "When I was using my laptop, I would get like saran wrap or something, and I kind of wrap it over the keyboard. So when I was pressing the buttons, I know it was kind of covered. So I could still press the buttons, but then the Saran wrap would be covering it."

To navigate digital information (e.g. recipes), 17 out of 20 participants mentioned using built-in gestures from screen readers of both laptops and smartphones to scroll through recipes. Our participants expressed dissatisfaction about **repetitive readings from screen readers** that they encountered frequently. P6 gives an example of when this occurred:

> "When I'm looking at the steps, I would have to listen to the steps as if in a situation where a step ended up, being like a whole paragraph of things you were supposed to do. I would listen to it. And then, you know, go back and complete a couple of

actions, and then go back and have to re-listen to the step from...you know it often plays from the beginning of the paragraph. If it's one big run-on sentence, then I have to re-listen to it from the start to see what the next ingredient is that I'm adding. And every time I go back, I end up having to listen to information that isn't, you know, necessarily relevant to the step I'm actually doing."

To acquire the information that they want to find, our participants also leveraged the search function of screen readers to **search for the header of different steps (e.g., step 5) or keywords (e.g., ingredient name, measurement unit)**. P1 explains how they use the search function:

> "I have found that, as you say, a lot of recipe webpages, they'll put their whole life story at the top, but I have found there is some consistency in those website designs. There's usually an ingredient heading on the website that denotes where the real recipe begins. Sometimes the design is not like that and I'll have to do a Ctrl-F and find and I'll search for a word of ingredients. or I'll search for the word "cup," because I know that's going to be in there somewhere. That is often why I'll go ahead and copy it into Notepad, because especially on mobile, it's a different experience. Ads on websites are different on mobile and I don't want to deal with that formatting in real-time. I'd rather do it when I'm sitting down and have the time to copy it out."

Regarding using tactile-based audio players and YouTube recipes, our participants **leverage conventional interactions**, such as play, pause, rewind, and fast-forward. P2 describes how he uses his Victor Stream, a handheld media player with the ability to browse the web [10]:

> "I'd have it on my Victor Stream, and I could just press the play/pause button and rewind it if I needed to, or fast-forward it."

Interestingly, we found that many of our participants dislike using smart speakers for navigating online recipes, due to the **lack of ability to find exact information, and difficulty the smart speaker has in tracking steps and reading status** (P3, P15, I1, I2). I2 describes his struggles with using the Amazon Alexa smart speaker:

> "Alexa can obviously have a lot of improvements made so that it's better able to guide you through these recipes. So like I said, if you're not using the correct language with Alexa, sometimes it will almost kick you out of the recipe and you entirely lose where you're at."

In our interviews, participants also highlighted the challenge of referring back to a recipe for missed information or details, and the heavy burden of memorization resulting from this struggle. This was a common issue that all of our participants faced, regardless of their preferred method for accessing information. One participant (P7) shared an example of when she struggled with both referring back to a recipe multiple times and memorizing information:

> "And I look, kind of skim back and forth, and try to keep remembering. 'Okay. So this much water for the rice. And no, this time when I make it, I'm not putting ginger in there, but I did the last time for the slightly other recipes that they sent me.' So yeah, looking back and forth to figure out what am I throwing in there, and when, and if my timing is right. So yeah, just scrolling back and forth as I went, as I would do steps."

Figure 4.3: WayAround Tagging System: Customized RFID tagging for recognizing different spice jars.

4.1.5 Information Identification

During the interviews, we asked our participants about their preferred approaches of identifying information during the cooking process. We found that our participants leverage crowd-based vision apps (e.g., AIRA, BeMyEyes), AI-based vision apps (e.g., SeeingAI), and customized technology (e.g., WayAround Tagging System (Figure 4.3)) for information identification purposes. From their responses, we also learned that our participants mostly leverage vision apps to measure ingredients, check on the status of food (e.g. doneness, visual characteristics), distinguish between objects that feel the same (e.g. a can of chickpeas versus a can of corn), operate kitchen appliances, and understand visual information on food packages.

Specifically, our participants emphasized the high frequency with which they use vision-based apps to recognize information on packages of food, such as identifying the food type, expiration date, nutritional information, and recommended cooking steps of the food (e.g., instant noodles). This demonstrates how key information when cooking using food in a package is not easily found through touch alone–vision is necessary to even find where text is located on a package of food, such as a box of cereal. P15 commented on the high amount of effort and time required when identifying food products:

> “So I use AIRA. I have used BeMyEyes...I think it’s discriminatory that blind people cannot just pick up a package and know what it is. And it’s frustrating to have to keep searching for the barcode. It’s frustrating to have to shake the can and think I know what it is. For instance, I think it’s green chilies and it ends up to be mushrooms. Or I open a package and it’s, I don’t know, something way different than I thought it was, almond milk instead of broth. That’s frustrating.”

From the information gathered from our interviews, we identified various challenges that our participants encountered while using vision apps in the kitchen. For example, they find it difficult to find text on food packages and point it in the field of view of a camera for a vision app to read aloud. They also struggle with how vision apps repeatedly and automatically read text in the camera’s field of view, which can be inaccurate, redundant, and frustrating. A third struggle they have is that vision apps need a high level of guidance from the user. P19 commented on how difficult vision apps, especially AI-based systems, are to use when recognizing information on food products:

> "Sometimes some of the scans, especially if you buy sweet condensed milk–If you buy canned fruit they might actually have a little recipe on the side of the can, and sometimes I try to read that and try to figure that out, and that's kind of an adventure– it's not very accurate at all times, but it's something I do just because I can't..."

Furthermore, our participants struggled with existing vision apps automatically reading and looping any text captured by the camera's lens, especially because it is annoying, unnecessary, and does not meet their navigation and information identification needs. For vision apps that offer live reading from the camera, even a slight movement of the camera can make vision apps stop in the middle of a paragraph and start over from the beginning. However, even vision apps that read photos, instead of text from a live camera, often fail to identify information due to other visual content (e.g. tables, text over background pictures). P9 recounts her struggles with using vision apps:

> "So [vision apps] those are not the best if you want to get the ingredients honestly, because it would just start reading whatever is in front of you, but there's no real order (...) So you cannot really explore what the device is saying. So, you cannot go backwards or forward within the text. It just starts talking and there's no way to stop it where you're interested at. If you want to know something in particular, it's just going to be talking and maybe if you got it in focus, great. But if you didn't, then it's hard to find. And with the document mode you take a picture, but most of the times packages have columns and tables. So, it's not very good at recognizing those kinds of things. So, it would be nice if there was a mode where you can really scan or take a picture of the product and it would give you a better interpretation of columns and better interpretation of tables."

Finally, our participants emphasized the importance of being able to restrict the type of information that vision apps provide and store, mostly for privacy and security purposes. P6 shares his concerns on the matter:

> "I don't want you to track other things. For example, I only want to know was it like a package of fish, or is it like a package of paddies. I only want to know this information while I'm turning on the system. But I don't. I don't wanna just like I don't want a system to predict what I want. This will reduce privacy concerns and also potentially be more accurate."

4.2 Information Needs of People with Vision Impairments in Cooking

We present the desired information that people with vision impairments want while cooking, as well as the structure in which they would like cooking information to be presented.

Recipe samples from the semi-structured interviews with inaccessible, visual descriptions
Bake until baklava is **golden and crisp**.
Drop the covered ice cream balls, into the hot oil and fry, turning occasionally so they **color evenly, until golden brown on all sides**.
Heat 2 tablespoons olive oil over high heat until **shimmering**.
Continue baking until **lightly browned on top**.
Fold the **sides, top, and bottom** of each husk in toward the center to enclose dough.
Add the onions to the skillet and sauté until **translucent**.
Slowly knead in a small amount of water at a time, using just enough to form a soft dough that is pliable and **even in color**.
Deep fry until **bright red in color** and crispy.
Reduce heat to 425 degrees F (220 degrees C) and continue baking until pastry is a rich, **golden brown**.

Table 4.2: Recipe samples from the semi-structured interviews with inaccessible, visual descriptions

Type of Information Content	Key Findings
Non-visual Interpretation	- Tactile and auditory descriptors should be provided, in addition to visual descriptors.
Summarization	- Recipes should begin with an overview, to avoid navigating through every recipe when trying to choose one in the Information Searching and Comparison Stage.
Precision	- Precise magnitudes of food status, quantities of ingredients, and sizes of tools should be provided.
Alternatives	- Substitutions for tools, ingredients, and cooking activities are helpful for BVI people.
Tolerance	- BVI people take longer to prepare ingredients than sighted individuals. When provided with exact cook times and when trying to account for the additional time used to prepare ingredients, food often ends up over- or under-cooked. - Time ranges, instead of exact cook times, should be provided.

Table 4.3: Information content that blind and visually-impaired people want during the cooking process. Our participants believe that making these changes to recipes would make them easier to preview and to execute.

4.2.1 Desired Information Content

We found that our participants prefer mainly just having an ingredient list (with quantities) and steps–they dislike having background stories, a recipe author's personal perceptions of how good a dish is, and shopping lists. In this section, we will present five main recipe structure accessibility preferences as a guideline for recipes designed for use by people with vision impairments.

Non-Visual Interpretation

In our interviews with participants, they gave detailed feedback on whether to add or remove any content from each step of the recipes we found. All of our participants pointed out the issues with having only visual descriptors in recipes, demanding other sensory descriptors. For example, Step 2 of one of the recipes used in P13's interview was: "Toast split bread under broiler. Remove bread when it is toasted golden brown in color." P13 expressed that "golden brown in color" is not useful information to blind people. In fact, all participants mentioned that recipes commonly only use visual descriptors to describe the status of food, which generates barriers for people with vision impairments to mentally convert visual descriptions to non-visual interpretations, such as using timing (P6, P7, P10, P11, P16, P19, I2), texture (P1, P2, P3, I2), sound (P1, P2, P13, P16, P18, P19) and smell (P4, P18, P19). Doing so not only requires more effort, but also requires the visually-impaired person to already have experience with cooking (e.g. what something golden-brown in color should feel like). An instructor, I2, suggested using timing and texture information to substitute visual indicators (e.g., golden brown, translucent):

> "I think as a society we're very visual. So a lot of recipes give you visual cues to tell you when things are either ready or where they should be. So like I said, the most prevalent one is, 'Cook until golden brown,' or the, 'Cook until the onions are translucent.' So I think that there could be additional information or that can support and help them understand what that means. So it's either, one, providing cooking times. So, 'Cook your onions for 10 minutes,' and that's when they'll be translucent. Or, 'Your onions should feel soft,' things of that nature. But because that's something either students have to typically experiment with or they have to Google it themselves. Cause like I said, a lot of information and recipes is they're giving visual cues, visual things to look for. So yeah, I think just providing other information or things that could tell you how to determine those things. And that's why I'm saying, a lot of students adapt and change the language so that they're better able to understand that."

Furthermore, across all non-visual indicators, we found that timing is more commonly used in recipes compared with other non-visual indicators (e.g. sound, aroma). However, we found that solely using timing cannot provide reliable outcomes, as cook times can vary from kitchen to kitchen. For example, the cook time of a steak may vary due to its thickness or the amount of heat a stove range provides. P19 explains why cook times alone are not enough:

> "So that's the thing. Sometimes if your water evaporates faster or if your heat is too low, it will be raw on the inside. So, it's not really telling you (...) if they're cooked through. Even after an hour passes, if you have your heat too low, then they might still be raw."

Summarization

Before starting to prepare ingredients, our participants usually make a full pass through a recipe to ensure they have all ingredients, essential tools, and enough time to make the dish. We found that it was very time consuming for our participants to going through entire recipes using a screen reader or another approach. During the reading process, they were forced to listen to and sometimes interact with an abundance of unimportant information (e.g. ads, clutter, pictures without alternate text). Therefore, 11 of our participants explicitly stated that they prefer having an overview of the recipe at the beginning to reduce the effort spent in choosing a recipe. P9 explains why a summary at the beginning would be helpful:

> "I also think getting a short overview of how the recipe is going to turn out, I guess, or some of the ingredients, maybe the basic ingredients that you're going to need without telling you the amounts necessarily. But at least that way you can figure out if, oh okay, I do have all the ingredients to make it, or oh, you know what, I'm missing something basic."

Beyond a high-level overview at the beginning of the recipe, our participants also mentioned the importance of having a short summary of the key tasks and estimated time required for a step, at the beginning of each step. Our participants explained that this would reduce the effort of navigating through a recipe. For example, P12 explains:

> "I would rather [the summary] be part of the steps, and they shouldn't go on too long. (...) It should be able to describe it pretty shortly. But I would like it to be right there in that step. You don't want to have to go to someplace else to find it, because then you'd have to find your way back to where you were."

Precision

Our participants requested that descriptions of ingredients and steps have more clarifications on magnitude, quantity, or size. Adding this information to recipes would reduce errors when cooking, because more non-visual options for knowing the status of food would be provided. For instance, in response to the recipe step "Form into patties and place on the prepared broiler pan or baking sheet," P16 noted:

> "I'd want to know an approximation of the size of the patty."

Beyond ingredients, we found that our participants want information about actions to be more precise. In response to the recipe step "Place foil around the edges of the crust to protect it from burning," P3 said:

> "Yep, how far on the edge? Is it over the edge into the pie? Is it the edge that is not covered by the egg mixture? I think I'd want a little more specific on that."

Furthermore, our participants expressed that the description of tools and appliances should be more precise. When we provided P9 with the recipe step "Heat oil in large skillet or wok over moderate heat," she said:

> "...The size of the skillet, because large could mean different things to different people..."

Alternatives

Our participants remarked that they often choose to use specific tools or cook in a certain way due to accessibility issues, such as inaccessible touchscreen-based appliances or actions that are highly visual (e.g., measuring). 12 of our participants expressed that they prefer recipes to contain substitutions for products, tools, and actions, as well as the ability to verify the validity of the substitutions. P9 explains the benefits of having substitution information for ingredients:

> "I guess it would also be cool to include in the recipe, maybe at the end or in the ingredients list, if there's any substitution for any ingredient that you can make if you don't have that ingredient at home, or if there's something that you can skip–optional things. I think mostly recipes do tell you if something is optional, but they don't give you any substitutions. For example, if you're trying to make cookies and then you don't have eggs, they don't necessarily give you substitutions for, okay, if you don't have eggs, we recommend you to use this...."

Regarding substitutions for tools and actions, P15 responds to the recipe step "Heat oil in a deep fryer to 400 degrees F (200 degrees C). Deep fry samosas until golden brown, 3 to 5 minutes each" by saying:

> "I don't deep fry things. What's an alternative? If there's a baking alternative, put it right there in the step, like: For deep frying, do this. For baking: do this. For air fryer: do this."

Tolerance

All of our participants experienced needing extra time to complete cooking steps non-visually, due to needing more time to find ingredients and tools in the kitchen, needing to operate devices with assistance from technology, and difficulties in measuring liquids. Due to the extra time needed to perform tasks in the kitchen, food is often either over-cooked or under-cooked (P5). P5 went on to recount his experience with getting sick eating under-cooked chicken: *"One time I cooked chicken, I probably didn't cook it long enough, and I got sick."* As a way to prevent this from happening, our participants expressed that they would like descriptions of cook time to provide a time range instead of an exact time, allowing for more tolerance for people with vision impairments. Along with the time range, a method of checking the food for doneness is necessary too. P4 comments on wanting a time range instead of an exact cook time:

> "Well, I know that they cannot give us the smell. But say five minutes timing and sometimes they can also say five to seven minutes, not just five minutes."

Beyond having tolerance on cook time, our participants also prefer having tolerance on the order of combining multiple ingredients together. In order to reduce the mental burden of memorizing the order of mixing ingredients, t should be stated clearly if the order matters or not. P11 describes having to memorize the order in which to add ingredients:

> "There is a lot of ingredients. I'm trying to remember what order they all are. I'm trying to remember what order they all go into directions."

Factor of Information Presentation and Structure	Key Findings
Information Density	- Recipes commonly found today contain steps that are too long. - Steps should be broken into multiple steps based on: (1) the introduction of a new ingredient or (2) a new action.
Ordering and Consistency	- Ingredients should be listed in order of use, to avoid confusion.
Grouping and Division	- Ingredients should be grouped based on type, separated by headers (e.g. dry ingredients, wet ingredients). - Steps should be ordered to optimize for time management. - In the ingredients list, quantities of ingredients should be provided for each dish component, not just overall quantities for the entire recipe.

Table 4.4: Information presentation and structure preferred by blind and visually-impaired people. Our participants suggested these improvements to how information is presented in recipes, and how recipes are structured.

4.2.2 Desired Presentation and Structure of Information

Information Density

In the interview, all participants mentioned that existing recipe steps are too long, and should be broken down into multiple shorter pieces for less effort in memorization and step following. Our participants prefer breaking the steps according to two main rules: the introduction of a new ingredient, and a new action (e.g. mix, chop). P19 explained how breaking steps down further based on the introduction of a new ingredient would lead to better information density:

> "When you were mixing the eggs white with the milk, that's one step. Then, when you fold in the egg yolks, I think that would be another step. That's another (...) agent entering. So anytime you have to add another ingredient, I would, but I consider that another step."

P11 gave an example of how he prefers breaking down steps based on new actions:

> "You have something, you're putting a bunch of stuff in and cooking it for 25 minutes, that's your first step. And then your next step is to put more stuff in, and cook that for six more minutes. The next step that changes or alters the cooking time or adds to the cooking time, makes that its own step. And then try to keep the activities in the step logically linked. So again, if you're manipulating food, cooking, adjusting heat, timing, that's one step."

Moreover, certain recipes contained similar actions multiple times in a single step (e.g., add milk whisk milk until combined, and then add milk and whisk again). Our participants wanted to break down repetitive actions into separate steps, so that they could better keep track of how many times a repetitive action was supposed to be done. P2 explains why he wants repetitive actions to be separated into different steps:

> "Maybe because it was adding, adding the milk, you know, and then whisking it until it thickened, and then adding more milk and whisking it until thickened. You might wanna break that down. So because you're adding milk. In fact, I can't remember if it was 3 or 4 times. You might want to add each one of those. So if I was doing something new, I'm, you know, might say, okay, I've done that step. Okay, now I'm on step 2, and so the second time I add the milk, and then you know, step 3..."

Ordering and Consistency

In addition to breaking up complex steps into multiple simple steps, our participants experienced confusion when ingredients within the ingredient list were not listed in order of use. They explained that it makes ingredient preparation/measurement and tracking ingredients difficult. P16 shared an example of how she prefers having ingredients listed in order of use:

> "I would say my ideal recipe is going to be my ingredients list first, grouped together based on usage. So, back to that banana bread, I want the baking soda, the flour, and the salt to be at the top of the list. And then I want it to be sugar, brown sugar, and bananas. And then I want it to be chocolate chips because those are the order that they're added in. And then I like to go into the directions section..."

Furthermore, our participants expressed that they spend a lot of time and effort when finding objects or gathering ingredients non-visually. To reduce the burden of doing these tasks for blind people, our participants want ingredients to be grouped based on ingredient type. For example, P3 explains how she wants baking ingredients to be grouped into two categories:

> "For ingredients, I'd want...A lot of times in baking recipes, wet ingredients together, dry ingredients together."

From our interviews with participants, we also identified inconsistencies in descriptions of ingredients between the ingredient list and the cooking instructions within the same recipe. Our participants expressed that these inconsistencies in descriptions caused confusion and often forced them to refer back to the ingredients list using screen readers on a laptop or verbal commands using a smart speaker. Navigating back and forth within a recipe is tedious and usually inaccurate, making following a recipe more difficult. P3 expresses confusion following an inconsistent description of an ingredient:

> "Are they dry? Is it dry garlic? Or was it garlic powder? Is it garlic salt? Whatever it is called for in the ingredient list, just call it by the same name all the way through, I'd say. Well, all the spices. Is it brown sugar? Is it white sugar? Is it ... I'm assuming it's ground cinnamon, and other things. Just keep consistency between the ingredient list and the step."

Grouping and Division

Based on the recipes we found when preparing for the interviews, as well as participant feedback, we found that existing recipes usually follow a sequential structure, which our participants execute in order from beginning to end. However, our participants complained that having a sequential order of steps is not time efficient in many scenarios, such as when the prep for one component of a dish can be done while another component cooks. In particular, because non-visual cooking takes longer than visual cooking (and what recipes say is the prep/cook time), time management is really important to blind cooks.

Regarding time management, our participants also expressed a preference in having separate sections for each component of food (e.g. steak, sauce, potatoes) with clear headers indicating the start of each section. These clear headers and sections would make navigating through a recipe using a screen reader or a smart speaker much easier, and prevent the device from getting lost, which our participants expressed often happens. P6 gives an example of when he believes this is essential:

> "...Making sure to you know if you're putting a whole dish together, and you're preparing a couple of different specific things that are gonna go together. Say you're cooking a sauce or a gravy, and then separately, you're cooking, you know, meat like a steak, or and then, you know, maybe you're also cooking (...) some onions and peppers to go on top of that. Clearly separating that into different sections is a big help for recipes."

Furthermore, if a certain ingredient (e.g. garlic) is used in multiple components of a dish (e.g. sauce, soup), our participants want it to be clearly noted in the ingredient list. Specifically, instead of just having the total quantity of the ingredient across all components, they would like the quantities of the ingredient needed for each component, separated by headers. P7 explained:

> "So you want two of them [garlic cloves] in this part and then one of them in that part. If more ingredients were that way, six cloves of garlic divided and then you'd know, Okay, let me just throw these in two separate areas so that I don't accidentally put all of it in one space..."

4.3 Design Features of Technologies to Support Information Access by People with Vision Impairments

Toward the goal of finding out how technologies can better support blind people in cooking, we asked our participants about their experiences with leveraging different ways to access information while cooking (e.g., AIRA/BeMyEyes, YouTube/podcasts, other people). We then used these scenarios as probes to further uncover the design features of technologies for people with vision impairments to access information throughout 5 different cooking stages. In this section, we present the following six design features: information granularity, guidance and feedback, status and timing, interaction methods, form factor and modality, and customization and adaptation.

4.3.1 Information Granularity

Based on our participants' responses, we learned that blind people prefer different levels of information, depending on their own cooking level and amount of experience. Unexperienced

Design Feature	Key Findings
Information Granularity	- Adaptive reading content should be provided, based on a BVI person's preferences and level of cooking experience. - Adaptive reading modes (entire recipe, step-by-step) can reduce the effort of navigation.
Guidance and Feedback	- BVI people who cook ask many contextual, recipe-specific questions that they would like answered in real-time. - Specialized guidance from blind cooking instructors and professional chefs is preferred.
Status and Timing	- Time management tools to help time simultaneous tasks can increase efficiency in the kitchen.
Interaction Methods	- Zero-touch interactions are preferred to avoid getting food on devices.
Form Factor and Modality	- A ubiquitous system of smart devices in the home would give BVI people options for which devices to interact with, and the modality of interaction.
Customization and Adaptation	- Customizable recipe editing tools in various modalities will decrease future effort of finding and saving recipes.

Table 4.5: We identified 6 design features that technologies should adopt in order to support blind and visually-impaired people when accessing information, and present ideas/suggestions for each design feature.

individuals often wanted detailed explanations of preparation steps (e.g., preparation of the toast of the bread) or instructions on how to use certain tools non-visually, while more experienced individuals expressed that this information would be annoying to listen to when reading a recipe. To offer flexibility in how detailed recipe information is, we propose **adaptive recipe reading based on a user's skills and preferences**. P9 explains that, with her culinary background, having less information is still manageable; however, in general, having too little or too much information are both undesirable:

> "It has to be detailed, but not so much that you have a paragraph for each step. So I don't want to get all the information, but I do want to get the information that I need. And my background is in culinary arts, so maybe for me, it's easier to understand even with fewer words. But if I learned to cook or started cooking while I was blind, I would definitely want to get some non-visual pointers in my recipes."

Furthermore, our participants expressed a desire for **adaptive reading modes while following the recipe during cooking (e.g., entire recipe, step-by-step)**. P1 provides a scenario in which she believes reading an entire recipe would be overwhelming for a blind cook:

> "I'm thinking...because I think there are different modes of viewing a recipe, so it could be that when someone is just looking at the recipe or when they're battle planning, they're preparing to cook, that information is annoying, it's too much. It is too much, it's cognitive overload to read all that..."

We discovered that our participants would like audio devices to use different voices/tones to classify different content information for clarity purposes (e.g., header, body of steps) (P12). Many participants expressed that they would like to be able to adjust the speed and volume of reading, to cater to different environments and cooking stages (P2, P3, P4, P14, P17, P20).

4.3.2 Guidance and Feedback

In our interviews with instructors, they recalled that their students constantly ask questions, mostly to confirm if what they are doing is correct. Often, instructors answer the questions verbally without hand-over-hand assistance (I1, I2, I3, I4) in order to foster independence. Our participants mentioned that existing ways of searching for recipes using screen readers, smart speakers, or vision assistance apps either cannot give them an expected answer or the process of finding an answer they are happy with is cumbersome and time-consuming. We found that people with visual impairments ask questions for confirmation throughout the cooking process as a way to reduce the effort in memorizing details. Being able to quickly and frequently get feedback reduces learning barriers for people with vision impairments during cooking. I3 explains why:

> "A lot of times it's just looking for reassurance. Did I do this right? That's the biggest one. I'm not sure if I'm doing this right or if this is the right piece of equipment. Because we try to get that out of the way ahead of time. I try to make them think through the process and ask most of the questions that they would've asked in the process ahead of time in conversation. It takes a lot of time to do that, but it saves time during the actual cooking process."

Our participants (I1-I4) remarked that cooking-related questions are often concise, contextual, and relevant to specific recipes or steps. For example, a blind cooking course student might ask “how long should I cook this tomato for again?” Our instructors’ responses highlight the importance of understanding a question’s intent in order to better guide blind people to cook. P3 further commented on why having questions answered in real time is helpful, saying:

> “Yeah. Standard questions like, what’s the cook time? What’s the oven temperature? Things like that. That would be handy to know right away.”

We also found that our participants prefer having specialized assistance, rather than receiving help from random people (e.g. crowdsourcing workers). They showed a preference for receiving professional feedback for cooking non-visually (e.g., describing actions non-visually). P11 expressed a wish for BeMyEyes to enable talking to chefs or other professionals:

> “So if they had some kind of connection to chefs or anyone that can note that they are cooks or they have good cooking knowledge, then when you go in and you specifically request: Hey, I need some help in the kitchen, then you can get someone who’s knowledgeable about cooking methods or just things that you need a little bit of guidance or things that you need to do something in the kitchen.”

4.3.3 Status and Timing

We discovered, based on our interviews, that our participants find it difficult to track steps and manage timing different dish components at the same time. I3 observed that the biggest change his students face is time management:

> “I think the challenge is time management. Time slips away from you real quick when it’s new, when they’re learning a new process. I do often have to remind folks, you can only have that hot for about a minute before you have to start the next process. So don’t fool around, just get it done and go to the next step or you’re going to burn what you’re after and your outcome will be different. Again, if it’s somebody that normally has problems with timing, I’ll let them burn a few things and let them see that they have to keep up on things.”

Time management is even more difficult when recipes require the cook to have multiple tasks running simultaneously (e.g., boiling pasta while preparing pasta sauce). P11 expresses his need for time management tools while performing multiple tasks at the same time:

> “I need time, and things that keep time management going. So a recipe where it will give you the instructions to cook pasta and tell you while you’re cooking the pasta now you can do something, you can make the sauce. You can do something else within the time frame, like the 11 minutes it takes to cook pasta.”

To ensure a device being used to read a recipe tracks the steps as a cook is reading or executing them, our participants are enthusiastic about the approach of interacting with the device to keep it up to date with the current reading progress. As for the interactions used, they suggested using voice commands and tapping on a device. P9 describes how she envisions interacting with a device to help it track her progress:

> “So how it works is I’ll ask for a recipe, it will give me the first option. If I like it, I can give it two different commands. The first command is start with ingredients or start recipe or something like that. I’m not getting the right wording, but it’s the

concept of. So when you start gathering the ingredients, it will go one at a time and you just tell like next ingredient. Of course you have to say her name and then next ingredient, next ingredient. And then it will tell you when you're at the end of the ingredients and it will ask you if you want to continue with the instructions and then it goes one step at a time. So, it's cool because sometimes one step takes a long time, like boil for 20 minutes. So, the recipe will still be there after the 20 minutes and you can ask what's next."

4.3.4 Interaction Methods

In the previous section, we presented that our participants rely heavily on sensations of touch obtained from touching food with their hands. This reliance is prevalent throughout all five cooking stages we identified. Because our participants' hands are often soiled with food, it is inconvenient for them to use their hands when interacting with smartphones, audio players, and laptops while cooking. All of our participants were of the opinion that having **zero-touch interactions** would reduce the time and effort necessary for people with vision impairments to interact with technology (and recipes) while cooking. P19 explained why he prefers hands-free interactions with devices:

> "It would be nice to have either something that can either be on some sort of a stand or hands-free something that you can ask your device to do, especially when you got your hands full, you know you're sitting there, and you that you're cooking a lot of the Times as well most of the time it's about timing right? So you got your potatoes in one thing, and you got your stuff, and it's really difficult to actually grab your phone and try to see? Okay, is this how much on this timer? How much it's left on this before it cooks? and then you gotta quickly drop it and go. You know. Raise your hands and then get back to other stuff. It'd be really nice to have something hands-free that can do that or even a wearable that concentrates on stuff like that cooking."

Our participants noted that cooking is a scenario in which mistakes can have a big impact on their overall experience. Because of this, participants preferred having restricted gestures to prevent incorrectly recognizing objects, using devices, or reading instructions. In support of having restricted gestures, P1 explained:

> "So sometimes with apps, you can restrict the gestures that are available to make so that (...) when you're navigating with a screen reader on the phone, you have all the gestures available to you so you can accidentally go out of the file or you can accidentally lose your place. So if it was an app, maybe a lot of these apps will turn off voice-over and have a self-voiced option, an app, that will restrict the gestures and only leave you a limited number of gestures available in that app. So you can't accidentally go out of the recipe."

4.3.5 Form Factor and Modality

When interacting with smart devices and other assistive technologies, the form factor and modality of the device is especially important; they can have a strong impact on the accessibility and usability of a device. Our participants emphasized the importance of having **ubiquitous systems that allow blind people to use in different environments** (e.g. a kitchen in the rehab center or a friend's kitchen) to ensure a similar setup and increased confidence during social activities due to more familiarity with the system. I1 emphasized how her students would benefit from having a ubiquitous system of smart devices:

> "Yeah, a huge, huge thing that I would love to support is having as many smart devices in the house. Although if a person is in training and they have the smart devices in the kitchen and they don't have them at their house, then that's going to be a challenge. So I also think meeting the student where they're at is a good thing, too. So meaning that besides the smart devices, if they don't have them at home, we'll then also teach them without those smart devices."

In addition to ensuring the devices are part of a ubiquitous system, our participants highlighted the importance of having connections and integration across different devices. This would reduce the reliance on a single device, making cooking using a device more fail-proof, in case one device is not convenient to see or interact with. P12 commented on his preference of having notifications on both phone and smartwatch: *"And then it'll notify me on my phone and my watch, my Apple watch when it is done."* I4, an instructor, also gave an example of when this would be useful:

> "I will ask her [Amazon Alexa] to look up a particular recipe. I don't know what it might be. Tuscan chicken for instance. And she'll have a whole list of recipes and she'll name them. And as we go through, she'll ask me if I want to gather ingredients or do I want to, I want her to send them to my phone. If I send them to my phone, then I just read them from the phone. But I have an A lady [Alexa] that I can move the speaker anywhere I want. So sometimes I'll just take the speaker into the kitchen and work on the recipe as she's giving me information."

4.3.6 Customization and Adaptation

In Section 4.1.2, we explained how our participants manually save recipes and edit them in Word documents or plain text files. Our participants frequently brought up the importance of having the ability to modify and customize the recipe. P8 expresses his practice of personalizing recipes after the first time cooking it:

> "My first time when following the recipe, I just get to follow the recipe. And wherever I feel, okay, I'm not comfortable with this, I can actually make my own personal adjustments. Yeah. Yeah. That's just my own recipe..."

Being able to save a personalized version of a recipe might be useful to cater to personal tastes, for example in the scenario that P5 provided:

> "I like garlic though. For example, it might say put a certain amount of garlic, I might want to add a little more."

Our findings show a need in having **customizable recipe editing tools for different modalities** (e.g., screen reader, smart speaker, wearable devices) (P5). Additionally, it is important for systems to provide intuitive ways for blind people to share and search for edits from other blind peers who also desire agency and personalization.

Chapter 5

Discussion and Future Work

In Chapter 4, we presented current methods, strategies, and challenges of accessing information during different cooking stages, the information need for people with vision impairments while cooking, and design features for technologies to support people with vision impairments in cooking. In this section, we will discuss our findings in-depth, categorized into: augmenting recipe information, augmenting information extraction, and augmenting interaction with information.

5.1 Augmenting Recipe Information

5.1.1 Non-visual Database for Common Cooking References

In our study, we discovered that our participants prefer having non-visual descriptors (e.g. cook until crusty to the touch) to supplement visual information (e.g., cook until golden brown). This was mainly to reduce learning barriers and contribute to increased inclusiveness and accessibility of online recipes. We learned that people with vision impairments either gradually learn to mentally convert visual descriptors into non-visual cues based on trial and error, or directly learn the corresponding non-visual cues from the instructors teaching blind cooking classes. The correct non-visual cues that correspond to visual descriptors may vary based on the ingredients, cooking method, size of food, shape of food, tools, and more. All of this leads us to believe that it would be beneficial to create a non-visual database of recipe terms and descriptions. Our proposed format is: [Original Visual Description], [Ingredient], [Cuisine], [Dish Name], [Cooking Method], and [Non-Visual Description]. We believe this database would reduce the learning curve for people with vision impairments when starting out with cooking. Furthermore, having such a database could contribute to future models that automatically transform recipes with visual descriptors to ones with non-visual descriptors as well. It would be interesting for future research to investigate the best interactions and navigation methods for this database.

5.2 Augmenting Information Extraction

5.2.1 Extraction and Conversion of Visual Content to a Non-visual Form

Based on the recipes we encountered while searching for recipes in our interview preparation, as well as our participant responses, we discovered that online recipes contain a plethora of distracting, inaccessible visual content that makes navigation difficult for blind people. For example, our participants complained about frequent pictures (without alternate text) with text to show the amounts of ingredients, or to demonstrate actions within a cooking step. As for video recipes, all of our participants said that they are not useful to actually cook along with, due to their precise visual and imprecise verbal descriptions of what to do. For example, in a video recipe, a person might verbally say "put the corn here" while putting the corn in a bowl. Since they didn't say where to actually put the corn, this step can be very confusing for blind people. In similar cases, we can leverage object recognition and data extraction tools to detect: what events occurred in the video (e.g. corn was put into a bowl), any text in the video to indicate timing, and sounds generated in the cooking process (e.g., sizzling). Using these pieces of information, we can create a clearer picture of what is going on in a video recipe; this approach could potentially contribute to the aforementioned non-visual cooking database and provide non-visual references for blind people to leverage while following a video recipe.

Physical Design of Packaging for Information Access

From our study, we learned that our participants use vision apps to identify objects, read text on food packages, and operate appliances. In Section 4.1.5, We also showed various challenges they struggle with while using vision apps. Prior research by Lei et. al. [36] also explored package fetching and identification by people with visual impairments. We propose future research in investigating and designing a universal food packaging template so that blind people can read the information on food packages quicker and with less effort. Specifically, we suggest trying layouts featuring designated locations on the packaging for certain information (e.g., nutrition label, ingredients list, cooking instructions), as well as tactile readings, such as those researched by Ribeiro et. al. [43]. Due to the difficulties blind people have with using existing vision apps (e.g. AIRA/BeMyEyes, SeeingAI) to recognize and read text on packages (Section 4.1.5), we advise package designers to adopt a plain text structure, to avoid putting text on top of a background picture, and to provide a summary of important information in a designated location on the package.

5.3 Augmenting Interactions with Information

5.3.1 LLMs for Kitchen-AI Systems

In Sections 4.1.3 and 4.3.2, we explained that our participants commonly ask instructors in their blind cooking classes questions for reassurance that they are doing the right thing. They also ask for clarifications on specific steps within a recipe. Here, we see an opportunity to build Large Language Models (LLMs) into assistive technology, using it them answer questions that are highly contextual and recipe-specific. Recently, BeMyEyes has incorporated GPT-4 into its app for image-to-text interpretation and image-related Q&A [4]. We also believe that LLMs can potentially help blind people in identifying objects and the status of food (e.g. doneness) [38] through the use of multi-modal data (e.g. images, audio). Beyond understanding the feasibility and efficacy of using LLMs to support blind people in cooking, future researchers should also explore potential biases, assumptions, and ethical issues that may come with using LLMs in these scenarios.

5.3.2 Multimodal Sensing Opportunities in the Kitchen

Based on our interview results, our participants prefer want the option of having multiple devices (and modalities) in the kitchen to interact with recipes (e.g., searching up a recipe on the smartphone and listening to it on a connected smart speaker). They would like the freedom to choose the modality with which they interact with the system and receive output. We believe that there are opportunities to explore: 1) What should the available input and output modalities in a connected system of devices be? 2) What are the preferred gestures and interaction methods for these multi-modal systems within the kitchen? 3) What are the design considerations of having multi-modal systems in the kitchen for blind people (e.g., deployment)? 4) What are the broader benefits to multi-modal systems in the kitchen, beyond interacting with recipes?

5.4 Future Work

Our participants consisted of 20 blind people with cooking experience and 4 blind cooking course instructors from a single vision rehabilitation center. We found that even among the four instructors, people have different methods of teaching people with vision impairments how to cook. This makes sense, as blind individuals have different learning styles as well. We believe that it would be valuable for future research to conduct interviews with people who have different relationships to blind home cooks (e.g., parents, friends, etc.) to develop a deeper understanding of the challenges that blind people face when learning how to cook. Additionally, we can valuable insight into how to better instruct blind people in cooking, which can translate into better assistive technology.

Chapter 6

Conclusion

We presented the existing practices and technologies that people with vision impairments when cooking. We also dove deep into the challenges that they face when accessing information through 5 cooking stages. We made novel findings related to information searching, extraction, interaction, and identification. Then, we explored the information needs that blind people have while cooking (e.g. desired information content, desired information presentation and recipe structure). We also suggested design features for future assistive technologies to better deliver accessible information while cooking (e.g., information granularity, interaction methods). We also created a set of design guidelines for future recipes to be less reliant on visual explanations, and to cater to blind people as well. In our Discussion, we also suggested future directions of research, including novel ideas of building LLMs into kitchen-AI systems, and multi-modal opportunities to support agency and reduce barriers to independent cooking. Overall, our paper provides a road-map to support information accessibility in cooking for people with vision impairments.

Bibliography

[1] (51) youtube. https://www.youtube.com/. (Accessed on 05/11/2023). 2.2

[2] Allrecipes — recipes, how-tos, videos and more. https://www.allrecipes.com/. (Accessed on 04/18/2023). 1, 2.2

[3] Amazon alexa. https://alexa.amazon.com/. (Accessed on 05/11/2023). 2.2

[4] Introducing our virtual volunteer tool for people who are blind or have low vision, powered by openai's gpt-4. https://www.bemyeyes.com/blog/introducing-be-my-eyes-virtual-volunteer. (Accessed on 07/04/2023). 5.3.1

[5] Bow tie pasta salad recipe (with tomatoes) — kitchn. https://www.thekitchn.com/bow-tie-pasta-salad-recipe-23533554. (Accessed on 05/11/2023). 2.2

[6] Introducing chatgpt. https://openai.com/blog/chatgpt. (Accessed on 05/11/2023). 2.4

[7] Jaws® – freedom scientific. https://www.freedomscientific.com/products/software/jaws/. (Accessed on 02/05/2023). 1

[8] Kroger chefbot – the kroger co. https://www.thekrogerco.com/kroger-chefbot/. (Accessed on 05/11/2023). 2.4

[9] Recipes, dinners and easy meal ideas — food network. https://www.foodnetwork.com/recipes. (Accessed on 04/18/2023). 1, 2.2

[10] Humanware — the all new victor reader stream. https://www.humanware.com/microsite/stream/index.html. (Accessed on 07/01/2023). 4.1.4

[11] The average number of ads on a google search recipe? 8.7. https://www.surgehq.ai/blog/the-average-number-of-ads-on-a-google-search-recipe-8-7. (Accessed on 07/09/2023). (document), 4.1

[12] Ali Abdolrahmani, Ravi Kuber, and Stacy M Branham. " siri talks at you" an empirical investigation of voice-activated personal assistant (vapa) usage by individuals who are blind. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 249–258, 2018. 2.3

[13] Niran Adetoro. *INFORMATION ACCESS FOR THE VISUALLY IMPAIRED IN THE DIGITAL AGE*. 10 2016. 2.2

[14] Riku Arakawa, Hiromu Yakura, Vimal Mollyn, Suzanne Nie, Emma Russell, Dustin P. De-

Meo, Haarika A. Reddy, Alexander K. Maytin, Bryan T. Carroll, Jill Fain Lehman, and Mayank Goel. Prism-tracker: A framework for multimodal procedure tracking using wearable sensors and state transition information with user-driven handling of errors and uncertainty. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(4), jan 2023. doi: 10.1145/3569504. URL `https://doi.org/10.1145/3569504`. 2.3

[15] Shiri Azenkot, Kyle Rector, Richard Ladner, and Jacob Wobbrock. Passchords: secure multi-touch authentication for blind people. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, pages 159–166, 2012. 2.3

[16] Jeffrey P Bigham, Craig M Prince, and Richard E Ladner. Webanywhere: a screen reader on-the-go. In *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*, pages 73–82, 2008. 1

[17] Marie Claire Bilyk, Jessica M Sontrop, Gwen E Chapman, Susan I Barr, and Linda Mamer. Food experiences and eating patterns of visually impaired and blind people. *Canadian Journal of Dietetic practice and research*, 70(1):13–18, 2009. 1, 2.1, 2.2

[18] Paul Blenkhorn. Producing a text-to-speech synthesizer for use by blind people. *Extra-ordinary Human-Computer Interaction: Interfaces for Users with Disabilities, ADN Edwards (ed.), Cambridge University Press, New York*, pages 307–314, 1995. 2.2

[19] Andrea Borghini. What is a recipe? *Journal of Agricultural and Environmental Ethics*, 28: 719–738, 2015. 2.4

[20] Stacy M Branham and Antony Rishin Mukkath Roy. Reading between the guidelines: How commercial voice assistant guidelines hinder accessibility for blind users. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, pages 446–458, 2019. 2.3

[21] Minsuk Chang, Léonore V Guillain, Hyeungshik Jung, Vivian M Hare, Juho Kim, and Maneesh Agrawala. Recipescape: An interactive tool for analyzing cooking instructions at scale. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–12, 2018. 1, 2.4

[22] Kathy Charmaz. *Constructing grounded theory: A practical guide through qualitative analysis*. sage, 2006. 3.3

[23] Giovanni Fusco, Ender Tekin, Richard E Ladner, and James M Coughlan. Using computer vision to access appliance displays. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, pages 281–282, 2014. 2.3

[24] Anhong Guo, Xiang'Anthony' Chen, Haoran Qi, Samuel White, Suman Ghosh, Chieko Asakawa, and Jeffrey P Bigham. Vizlens: A robust and interactive screen reader for interfaces in the real world. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 651–664, 2016. 2.3

[25] Anhong Guo, Jeeeun Kim, Xiang'Anthony' Chen, Tom Yeh, Scott E Hudson, Jennifer Mankoff, and Jeffrey P Bigham. Facade: Auto-generating tactile interfaces to appliances. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5826–5838, 2017. 2.3

[26] Rex Hartson and Pardha S Pyla. *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier, 2012. 3.3

[27] Liang He, Zijian Wan, Leah Findlater, and Jon E Froehlich. Tactile: a preliminary toolchain for creating accessible graphics with 3d-printed overlays and auditory annotations. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 397–398, 2017. 2.3

[28] Apple Inc. Accessibility - vision - apple (ca). `https://www.apple.com/ca/accessibility/vision/`, March 2021. (Accessed on 03/16/2021). 2.2

[29] Zoom Video Communications Inc. Video conferencing, cloud phone, webinars, chat, virtual events — zoom. `https://zoom.us/`, September 2012. (Accessed on 07/20/2021). 3.3

[30] Nabila Jones, Hannah Elizabeth Bartlett, and Richard Cooke. An analysis of the impact of visual impairment on activities of daily living and vision-related quality of life in a visually impaired adult population. *British Journal of Visual Impairment*, 37(1):50–63, 2019. 1, 2.1

[31] Shaun K Kane, Jeffrey P Bigham, and Jacob O Wobbrock. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, pages 73–80, 2008. 2.3

[32] Fumihiro Kato and Shoichi Hasegawa. Interactive cooking simulator: Showing food ingredients appearance changes in frying pan cooking. In *Proceedings of the 5th international workshop on Multimedia for cooking & eating activities*, pages 33–38, 2013. 2.3

[33] Andrey Khusid and Oleg Shardin. An online visual collaboration platform for teamwork — miro. `https://miro.com/`, 2011. (Accessed on 06/02/2021). 3.3

[34] Eliza Kostyra, Sylwia Żakowska-Biemans, Katarzyna Śniegocka, and Anna Piotrowska. Food shopping, sensory determinants of food choice and meal preparation by visually impaired people. obstacles and expectations in daily food experiences. *Appetite*, 113:14–22, 2017. 1, 2.1

[35] Kazuma Kusu, Nozomi Makino, Takamitsu Shioi, and Kenji Hatano. Calculating cooking recipe's difficulty based on cooking activities. In *Proceedings of the 9th Workshop on Multimedia for Cooking and Eating Activities in conjunction with The 2017 International Joint Conference on Artificial Intelligence*, pages 19–24, 2017. 2.3, 2.4

[36] Wentao Lei, Mingming Fan, and Juliann Thang. "i shake the package to check if it's mine" a study of package fetching practices and challenges of blind and low vision people in china. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2022. 5.2.1

[37] Barbara Leporini, Maria Claudia Buzzi, and Marina Buzzi. Interacting with mobile devices via voiceover: usability and accessibility issues. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, pages 339–348, 2012. 2.2

[38] Franklin Mingzhe Li, Jamie Dorst, Peter Cederberg, and Patrick Carrington. Non-visual cooking: exploring practices and challenges of meal preparation by people with visual

impairments. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–11, 2021. 1, 2.1, 4.1.4, 5.3.1

[39] Mingzhe Li, Mingming Fan, and Khai N Truong. Braillesketch: A gesture-based text input method for people with visual impairments. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 12–21, 2017. 2.3

[40] Google LLC. Get started on android with talkback - android accessibility help. `https://support.google.com/accessibility/android/answer/6283677?hl=en`, March 2021. (Accessed on 03/16/2021). 2.2

[41] Weiqing Min, Shuqiang Jiang, Linhu Liu, Yong Rui, and Ramesh Jain. A survey on food computing. *ACM Computing Surveys (CSUR)*, 52(5):1–36, 2019. 1

[42] Tim Morris, Paul Blenkhorn, Luke Crossey, Quang Ngo, Martin Ross, David Werner, and Christina Wong. Clearspeech: A display reader for the visually handicapped. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(4):492–500, 2006. 2.3

[43] Gisele Yumi Arabori Ribeiro, Maria Lilian de Araújo Barbosa, Maria Lúcia Leite Ribeiro Okimoto, and Rafael Lima Vieira. Information for tactile reading: A study of tactile ergonomics of packaging for blind people. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018) Volume VII: Ergonomics in Design, Design for All, Activity Theories for Work Analysis and Design, Affective Design 20*, pages 1682–1688. Springer, 2019. 5.2.1

[44] André Rodrigues, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. Getting smartphones to talkback: Understanding the smartphone adoption process of blind users. In *Proceedings of the 17th international acm sigaccess conference on computers & accessibility*, pages 23–32, 2015. 2.2

[45] Alexander Russomanno, Sile O'Modhrain, R Brent Gillespie, and Matthew WM Rodger. Refreshing refreshable braille displays. *IEEE transactions on haptics*, 8(3):287–297, 2015. 2.2

[46] Ayaka Sato, Keita Watanabe, and Jun Rekimoto. Mimicook: a cooking assistant system with situated guidance. In *Proceedings of the 8th international conference on tangible, embedded and embodied interaction*, pages 121–124, 2014. 2.3

[47] Yayoi Shimomura, Ebba Thora Hvannberg, and Hjalmtyr Hafsteinsson. Accessibility of audio and tactile interfaces for young blind people performing everyday tasks. *Universal Access in the Information Society*, 9:297–310, 2010. 2.2

[48] Dan Tasse and Noah A Smith. Sour cream: Toward semantic processing of recipes. *Carnegie Mellon University, Pittsburgh, Tech. Rep. CMU-LTI-08-005*, 2008. 2.4

[49] Ender Tekin, James M Coughlan, and Huiying Shen. Real-time detection and reading of led/lcd displays for visually impaired persons. In *2011 IEEE Workshop on Applications of Computer Vision (WACV)*, pages 491–496. IEEE, 2011. 2.3

[50] ChunYuen Teng, Yu-Ru Lin, and Lada A. Adamic. Recipe recommendation using ingredient networks. *CoRR*, abs/1111.3919, 2011. URL `http://arxiv.org/abs/1111.3919`. 2.4

[51] Carlton Anne Cook Walker. Hands on? hands off! https://nfb.org/sites/default/files/images/nfb/publications/fr/fr34/1/fr340101.htm, 2015. (Accessed on 07/01/2023). 4.1.3

[52] S. YOSHIKAWA, S. NISHIMARU, T. TASHIRO, and M. YOSHIDA. Collection and classification of words for description of food texture. *Journal of Texture Studies*, 1(4):443–451, 1970. doi: https://doi.org/10.1111/j.1745-4603.1970.tb00743.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-4603.1970.tb00743.x. 2.1

[53] Massimiliano Zampini and Charles Spence. The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies*, 19:347 – 363, 02 2005. doi: 10.1111/j.1745-459x.2004.080403.x. 2.1