URBAN ECHOES

ENGAGING CITY DWELLERS IN THEIR LOCAL BIRD WILDLIFE USING AESTHETIC EXPERIENCES

JAKOB WORM, 202004405

MASTER'S THESIS

May 2025

Advisor: Peter Gall Krogh

AARHUS
UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

URBAN ECHOES

Engaging city dwellers in their local bird wildlife using aesthetic experiences

JAKOB WORM

Master's Thesis

Department of Computer Science
Faculty of Natural Sciences
Aarhus University

May 2025

ABSTRACT

Short summary of the contents in English...

Contents

1 Introduction 1
- 1.1 Generative AI Declaration 1
- 1.2 Urban Echoes: An Auditory Map Experience for Avian Engagement 1
  - 1.2.1 Application design 2
- 1.3 The role of technology in the preservation of local environments 2
- 1.4 Noticing as a design tool and nature appreciation 5
- 1.5 Using aesthetic auditory experiences for engaging people in local birdlife 7
- 1.6 Project timeline 8

2 Related Work 9
- 2.1 Engaging With Nature Through Audio Content 9
  - 2.1.1 Nature soundscapes: an audio augmented reality experience 9
  - 2.1.2 Audio-enabled locative media 10
- 2.2 Applications for engagement with birds 10
  - 2.2.1 Citizen science and bird observation 11
- 2.3 Digital storytelling 13
- 2.4 audio based interfaces 14
- 2.5 Design as research and drifting 14

3 Design analysis 15
- 3.1 A different direction: Creating a user-maintained bird dataset for visual recognition 15
- 3.2 Audio design 16
  - 3.2.1 The Challenge of Aligning the Acoustic Ecology with the Visual Environments 17
  - 3.2.2 Evaluating the sonic aesthetics 17
- 3.3 Designing the listener role 18
  - 3.3.1 Mapping the world 19
  - 3.3.2 Mapping sounds 19
- 3.4 Adding time temporal aspects Seasons 20
- 3.5 Creating a fully auditory application 20

4 Design and Evaluation of System Architecture 21
- 4.1 General system Architecture 21
- 4.2 Flutter and Dart 21
- 4.3 Vertical slice of the prototype 21
- 4.4 Voice recognition and audio 24

4.5 Backend 24
4.5.1 Database 24
4.6 Speech recognition, text to speech and recordings 26
4.6.1 Text to speech 27
4.7 Auditory considerations and how it translates to code 27
5 Implementation 28
5.1 Auditory map(location service) 29
5.1.1 bird_sound_player 29
5.1.2 background_audio_service 29
5.1.3 season_service 29
5.1.4 azure_storage_service 30
5.1.5 location_manager 31
5.1.6 How it all works together(Need more in depth rewrite now that I have explained all the dependencies) 32
5.2 Observation uploader and database connections 32
5.3 RecordingPlayerService 32
5.4 Speech coordinator service 32
5.5 Backend 32
6 Evaluation 33
6.1 Methodology 33
6.1.1 Participants 34
6.2 Results(This section is subject to change as all the evaluations are not finished yet) 34
6.2.1 Interview results 34
6.2.2 Results from the Dairy study 36
6.2.3 Final relections 41
6.3 Discussion (far from done) 42
7 Conclusion 43
7.1 Future work 43
7.2 FINITO 43

I Appendix
A Project Work on More than human-centred design 45
A.1 The report from the project work 45
A.2 The early prototype for Urban Echoes 45
B Github and code 46
B.1 Multi provider setup 46
B.2 Fastapi backend 48
.1 FastAPI Backend Code 48
A Evaluations 54
A.1 Diary study responses 54

Bibliography 55

List of Figures

Figure 1.1 Screenshots from the Urban Echoes application. 3
Figure 2.1 A screenshot of the eBird live map, the 31-03-2025[58] 12
Figure 3.1 Annotated image from the system description document for the initial idea 16
Figure 3.2 The definitions of the dimensions from the paper *Towards a Framework of Aesthetics in Sonic Interaction* [16] 18
Figure 4.1 The visuals for the vertical slice 23
Figure 4.2 Screenshots illustrating the speech recognition debug visuals in the application. 25
Figure 6.1 A quick sketch by participant 4 drawn after explaining her idea for bird neighbourhoods 39

List of Tables

Table 4.1 Schema for the `bird_observations` table 26
Table 4.2 Schema for the `birds` table 26

LISTINGS

5.1 MultiProvider setup in Flutter 28
5.2 Azure Storage Service Initialization Method 30
5.3 Dart Function: _handlePositionUpdate 31
B.1 MultiProvider setup in Flutter 46

2 Full FastAPI Backend Code 48

ACRONYMS

HCI Human Computer Interaction
SHCI Sustainable Human-Computer Interaction
SID Sustainable Interaction Design
SDG Sustainable Development Goals

1 INTRODUCTION

1.1 GENERATIVE AI DECLARATION

For this thesis, I have used Generative AI as a programming aid. The models used in order of magnitude are Claude 3.7 Sonnet, GPT3 and GPT4 in the form of GitHub CoPilot, and Chatgpt 4. All three have been used to provide suggestions and improvements to the codebase and for code generation and completion. GPT4 have also been used to create BibTeX citations for websites with prompts of a similar nature as the following prompt: *https://www.inaturalist.org/ Can you create a BibTeX citation for this website*. Lastly, I have used ChatGPT to do a translation for the Danish evaluation answers. After doing a translation, I read the resulting text and made sure the translation was as accurate as possible. Both the original and translated version can be found in Appendix A.1.

1.2 URBAN ECHOES: AN AUDITORY MAP EXPERIENCE FOR AVIAN ENGAGEMENT

In this thesis, I investigate an aesthetic auditory map experience called Urban Echoes and its potential to engage users in their local bird wildlife. Furthermore, I detail the application's creation, decisions, and underlying work, along with discussions on related work and how my work differs. Lastly, the thesis covers the evaluation and results thereof, including prospects for future work.

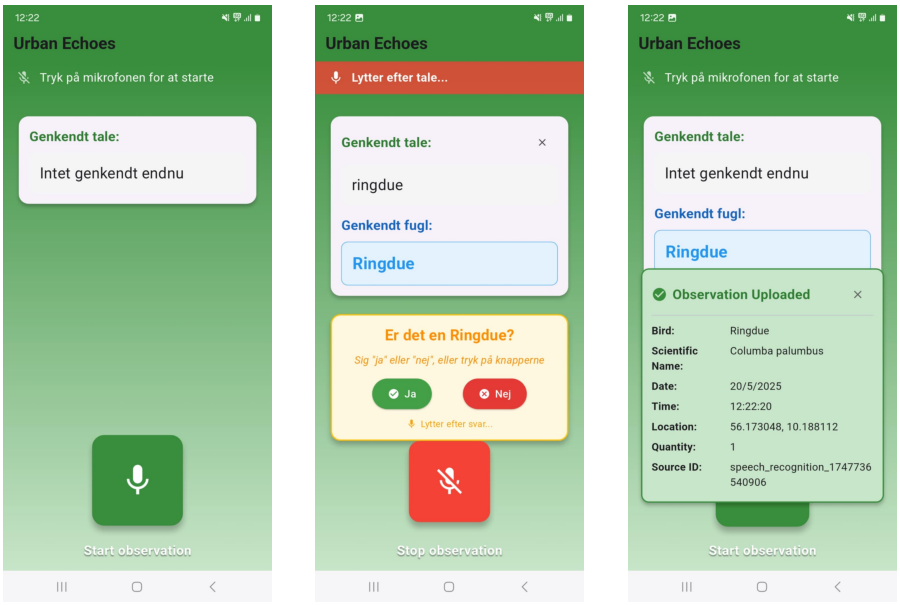
The idea for Urban Echoes is a continuation of a previous idea developed in a 10 ECTS project on more than human-centred design[15] that I conducted together with Magnus Lasse Lund Bentsen, a fellow IT product development master's student, last semester. The 10 ECTS project ended with a report and a rough prototype of this original idea, which can be found in Appendix A.1 and Appendix A.2, respectively. However, the continuation into a master's thesis has seen many alterations in the idea, as such, the original work is not required to understand the concept.

1.2.1 Application design

Urban Echoes consist of two primary activities, dubbed the observer and listener activity. Users can make sightings/observations by seeing or hearing birds in their immediate environment. The user can make a observation by clicking the applications microphone button as seen in figure 1.1a and speaking the observed birds name. This prompts the application to respond asking for a confirmation by presenting a confirmation card seen on Figure 1.1b along with asking the user if this was the bird they saw. If the user responds "ja"(yes) the observation will be saved and uploaded to a central database and a upload successful card seen in Figure 1.1c will be displayed along with a recorded message telling the user the observation have been successfully uploaded. The observation will consist of the bird observed featuring both common and scientific name, along with the geospatial location, date and time of day. If the application recognizes a different bird than the one the user said, the user can respond with "nej" (no). This prompts the application to ask if the user meant one of the following birds, presenting the most likely alternatives based on what it heard. This interaction is illustrated in Figure 1.1d. If the bird observed is not among the presented alternatives, the user can click "ingen af dem" (none of them). The app will then display a search bar where users can manually enter the bird they observed, as shown in Figure 1.1e. This is considered a last resort, as the application is primarily designed to be used through its auditory interface. The listener activity allows users to explore the auditory map created using the observations. A point will represent the each observation located at the position where it was recorded. The point will have a radius of 50 meters, with the playback audio increase in loudness the closer to the observation centre a listener might be. Observation points can overlap with up to five points playing simultaneously. The user can then move around the city to hear the observations made by other users. The application design will be described in more detail in Chapter 3.

1.3 THE ROLE OF TECHNOLOGY IN THE PRESERVATION OF LOCAL ENVIRONMENTS

The call for climate action and nature conservation has in recent years gained more attention, and rightly so. With the global community behind on 41 out of 42 indicators to achieve the 2030 goals of the Paris Agreement[7], the prospects of achieving these goals become more and more difficult.



(a) Start Screen **(b)** Confirmation Card **(c)** Observation Made

(d) "Was it one of these?" **(e)** Search Bar **(f)** Search Results

Figure 1.1: Screenshots from the Urban Echoes application.

Within Human-Computer Interaction Human Computer Interaction (HCI), the challenge for climate action has been seized, and the resulting fields, of Sustainable HCI Sustainable Human-Computer Interaction (SHCI) and Sustainable Interaction Design Sustainable Interaction Design (SID), were created after two influential papers by Eli Blevis [6, 33]. These new fields focused on using technology to promote sustainability in the user's life across many sectors, such as the environment, social justice, public health, and other areas concerning building a sustainable future.

A review of the literature by Hansson et al. (2021) surveying the field of SHCI found that research in this area is related to a wide range of the United Nations Sustainable Development Goals Sustainable Development Goals (SDG), with published work that addresses 6 of the 17 goals. This highlights the interdisciplinary and wide-reaching nature of the field [25]. Much of the research focuses on the direct application of technology to address specific SDG-related challenges. For example, Hansson et al. Hansson, Cerratto Pargman, and Pargman [25] report that of the 26 studies mapped to SDG Goal 12.2 - 'By 2030, achieve sustainable management and efficient use of natural resources' - the majority (18 articles) explored ecofeedback or ecovisualization systems. These systems aim to make resource consumption more visible and understandable and promote sustainable behaviours.

In contrast, the paper "Have We Taken On Too Much?: A Critical Review of the Sustainable HCI Landscape" by Bremer et al. argues that SHCI is not suited to create large-scale change across large and complex problem areas involving various disciplines with the direct application of technology. Instead, they argue SHCI researchers should seek to incorporate Green policy informatics[8] into their projects. Green policy informatics is the idea that technology should help promote climate action by providing tools for transparency and supporting complex decision-making. Beyond that, it should encourage communities and support the push for green policies.

I believe both approaches have their merits. My project is closer to the more direct application school of thought of SHCI as there are no direct pathways to incite policy changes or form communities, but rather a focus on inciting the individual's interest in local wildlife. This comes from a focus on individual interests and emotional experiences rather than actively pushing for radical changes or the creation of communities. This distinction is important and can be exemplified in how the prototype facilitates the creation of a collective artefact in the form of the auditory map. However, the focus is on the observations rather than the individual who recorded them. I will discuss the de-

sign choices in more detail in Chapter 4.

To justify my decision to prioritise the individual's emotions over community creation and focus primarily on those living in urban environments, I will point to the global development of increasing urbanisation. According to UN projections, 68% of the world's population is expected to live in urban settings by 2050[55]. As wild nature is more scarce in urban environments than in rural areas, it follows that contact with wild nature will, for an increasing number of people, be something to seek out instead of something they passively encounter. Exacerbating the problem of scarcity of wild nature in urban settings is the fact that key urban-dwelling animals who thrive in cities such as pigeons, raccoons, and rats, have poor reputations, leading to negative or indifferent meetings between urban dwellers and local wildlife[19]. Although many birds are commonly positively received, Urban Echoes aim to create memorable and positive experiences with all birds, even the commonly unpopular ones, such as gulls and pigeons.

1.4 NOTICING AS A DESIGN TOOL AND NATURE APPRECIATION

The act of *noticing*[54] as introduced by anthropologist Anna Tsing has gained traction within the SHCI community[32]. In this context, *noticing* extends to more than perceiving what is in front of us, but noticing the complex relationships surrounding our object of notice. This includes considering how economic, social, environmental, and other contextual factors shape and transform what is being observed. It calls for a heightened curiosity and a sensitivity on the part of the observer—an openness to perceiving more than what is readily apparent.

As designers, we can ask ourselves how the act of noticing can be used to inform more than human-centred design. As previously discussed in chapter 1.2, this project originally started as a 10 etcs exploration of the more-than-human centred design space. Although we were unaware of the concept of *noticing* during that project, we inadvertently did it to some degree during our attempts to understand the city pigeons, the focus of the study at the time. I stated the matter thesis with the knowledge and newfound appreciation for pigeons acquired during that project. As I got further into my master's thesis, my interest in pigeons expanded into birds in general, and I found myself noticing the birds in the city in ways I had never done before. I started perceiving them more, and noticed which birds were found in which environments.

need a brigde

Noticing as a way of increasing the designer's ability to decenter the human has been investigated in the text*Watching Myself Watching Birds: Abjection, Ecological Thinking, and Posthuman Design*[5] by Biggs et al, where the main author of the paper performed an auto-etnograpic study on how noticing the birds of Indiana changed her perceptions and led to growth both as a designer and person. In the paper, she describes how she develops feelings of abjection in the author's words, a simultaneous fascination and repulsion that arises towards the birds from her process. I did not share this feeling of abjection from noticing the birds, on the contrary, I felt a newfound appreciation and excitement from starting to recognise the birds and perceiving them more. That being said, my newfound focus on sonic environments also makes me pay more attention to the city's other sounds, with cars being unpleasant, whereas before I would more so consider background noise and pay it no mind. I want to make it clear that even though my experiences and feelings differed from Biggs' that neither is inherently more correct than the other.

The SHCI community can provide tools through interactive systems that help people notice in otherwise impossible ways, by giving us new senses, different perspectives, and more. Three examples can be found in Rosén et al's work, where they present three design concepts, each exemplifying different ways of noticing in a city gardening community[46]. A control-oriented approach to noticing, focusing on monitoring factors in the case of the gardening community, which included factors such as soil ph values, moisture, etc. The second approach to noticing they present is sensibility-oriented, focusing on developing tacit knowledge over time. Lastly, they present an appreciation approach focused on sensory engagement and exploratory experiences, without the need for practical application[46].

While urban echoes design does not force users to decenter humans, it encourages users to notice birds. Both my observations and existing literature [5, 32, 39, 46] suggest that the act of noticing is a powerful tool for fostering more-than-human-centred design. As designers—and people in general—deepen their appreciation and understanding of what they notice, they become more attuned to the needs and presence of non-human subjects. Urban Echoes are closely aligned with the appreciation approach of noticing. Providing users with the ability to hear the sounds in isolation, at least if they are using noise-cancelling headphones.

1.5 USING AESTHETIC AUDITORY EXPERIENCES FOR ENGAGING PEOPLE IN LOCAL BIRDLIFE

Urban Echoes is a collaboratively created audio map of local birdlife, designed to engage urban residents with the natural environment surrounding them. This project is grounded in the hypothesis that fostering participation through positive experiences can support conservation and reintegration efforts. This is supported by Chawla et al. [11] who found that our experiences with nature and connection to it play a significant role in shaping our willingness to act on climate change. Specifically, they found that formative experiences with wild nature are crucial in developing environmental sensitivity, which in their work is the motivation to learn about, care for, and feel concern for the environment[11]. To create experiences, the design for Urban Echoes is conceptually informed by the view of the aesthetics of interaction, as described by Petersen et al. [42]. In this framework, aesthetics are considered in terms of both context and use. This perspective recognises that while a design may have aesthetic potential in itself, the aesthetics of use emerge through the user's process of sense-making and the personal meanings they derive from the interaction. To illustrate how this view can be used to analyse academic work seeking to engage people in local nature, we can look at Gaver et al.'s Naturewatch Camera[24]. Naturewatch is a citizen science project in which users construct DIY wildlife cameras that autonomously take images of wildlife when within the frame. The aesthetics of use and context can be seen in how users evaluate the cameras, where sense-making and meaning become paramount to the user experience. Users comment on how the cameras have allowed them to observe the lives of the animals in their gardens without the animal recognising their presence, something they have previously been unable to do [24]. The community aspects and sharing of images also play into the aesthetics of the experience, as it's an important part of the context and meaning for the users.

Maybe I should reread and include making sense of things from saints

For my project, birds have been chosen as the subject for the application because of their prevalence worldwide, the diversity in species, and the perceived pleasant qualities of their soundscapes. Bird sounds have even been found to help with stress and attention recovery[44], showcasing the beneficial effect their presence can have on humans. It's important to note that the types of sounds that have these effects depend on personal preference and experience [44]. One of the aspects cited as restorative is the feeling of being connected with nature[44]—something I wish to test if Urban Echoes recreate despite digitalised audio. Digitalised sounds have been used in similar work by Lawton et al.[30] who investigated the use of speakers in a forested

area to play bird calls. They found that nature soundscapes had numerous beneficial effects even when digitalised, such as making people pay more attention to their visual surroundings, increasing their perceived enjoyment, and making them think about the loss of nature[30]. This hints towards the possibility of increased sonic awareness also raises visual awareness. The Urban Echoes application is geared towards urban dwellers, as urban soundscapes have considerably more noise pollution[43] than rural areas.

1.6 PROJECT TIMELINE

The idea and concept were expanded from the 10 etc project in the first weeks of the project. Development of the application took place from the end of January to early May 2025, with report writing happening in parallel to the coding, starting in mid-February until the delivery date in mid-June. Evaluations were held in May, starting on the 12 or the 14 of May; each evaluation took place a week later, with most being started on the 12 and the rest on the 14.

2 RELATED WORK

In this thesis section, I present related work that has inspired or reflected my project. Furthermore, I discuss similar commercial and scientific apps, their contributions and how Urban Echoes differs.

2.1 ENGAGING WITH NATURE THROUGH AUDIO CONTENT

Audio-based technologies are promising for fostering connections between humans, animals, and the environment. Several studies demonstrate how soundscapes and audio-augmented experiences can engage participants while raising awareness about ecological systems.

2.1.1 Nature soundscapes: an audio augmented reality experience

Lawton et al. demonstrate how augmented reality with audio has the potential to highlight the loss of nature[30] and make people reflect. To reintegrate animals and keep the ones we have, we need to be aware of the animals, how they adapt to humans and cities and what we can do to create spaces for all. Participants in their project became more aware of the loss of nature, with the soundscapes having the added benefit of fostering feelings of calmness and engagement with the local environment. However, certain trade-offs must be considered when implementing soundscape experiences. Although binaural audio through headphone playback is mentioned as an area of interest in Lawton's, it also has some limitations. As noted by Lawton et. al, "binaural audio, via headphone reproduction, detaches the listener from the real-world sounds in the environment" [30].

For any similar project like Urban Echoes, maintaining a connection to the immediate natural environment is critical to the augmented acoustic reality (AAR) experience. Encouraging participants to pay greater attention to the natural environment aligns with the objectives of Urban Echoes; however, in this project, the playback audio will be fully detached from the present physical environment and instead represent a recorded event. This partly stems from a desire to balance goals with ecological and practical considerations. There has been very little research into the effects of using playback audio of bird sounds in communal spaces, such as cities. However, one paper by Harris et al.

investigated the effects of playback audio on two tropical birds. They found it had multiple adverse effects, exposing the birds to danger and increasing stress levels[26]. Beyond this, multiple large birding organisations such as the American Birding Association, Rochester Birding Association, and more advise against using playback, citing problems such as disrupting nesting, breeding and the bird's ability to raise the young[2, 3, 48, 49].

Focusing on individual immersive experiences rather than communal sound installations. The design minimises ecological disturbance while allowing participants to focus on the presented audio content, albeit at the cost of their immediate auditory surroundings. The design tries to remedy this by including the observer role to encourage users to experience real-world sounds and create observations in content-less areas. The underlying hypothesis is that increased sonic awareness also increases the user's visual awareness of nature around themNeed to see if somebody has looked into this.

2.1.2 Audio-enabled locative media

Urban Echoes also draws inspiration from Audio-enabled locative media like audio walks. The first Audio walks were made somewhat serendipitously by Janet Cardiff in 1991[9]. This first work inspired her whole career, which features the creation of multiple acclaimed works. Her most famous audio walk *Her long black hair*[10]. The audio walk takes the listeners through Central Park South and uses pictures along with the audio and location to create the experience. Common for audio walks is the focus on geographical context, active participation in the form of walking to and from and a narrative structure.

"placed sound" (Behrendt 2012) "situated sound" (Fagerjord 2011)

2.2 APPLICATIONS FOR ENGAGEMENT WITH BIRDS

There are many free applications for identification and participation in local wildlife. The following is not an exhaustive list, but an exemplar-based description of the field. The most prominent app in the field is Merlin ID[40], developed by the Cornell Lab of Ornithology. The app allows users to identify birds based on their calls and songs. Furthermore, users can take images or describe the bird using pre-made queues and then get shown suggestions based on their answers. Looking at their app page, it's apparent that the app's goal is to engage people in birding and make the hobby more accessible to hobbyists and casual users. To ensure accuracy, MerlinID suggestions are based

on regional sightings made on eBird[51], one of the large citizen science[27] websites where birders can upload bird sightings.

Compared to Urban Echoes, MerlinID has a larger focus on the educational aspect of learning to recognise birds, helping users increase understanding and knowledge. Furthermore, Merlin ID encourages users to have an increased sonic awareness and listen for the birds in their immediate environment. Urban Echoes, on the other hand, is less concerned with natural sounds and the soundscape. Instead, it presents an exaggerated soundscape to the users. Personal reflections from the use of both apps showed clear strengths with each. With Urban Echoes, the immediate visual environment became the basis for where the bird could have been seen, prompting a reflection on which environments the birds are common in. I also found myself looking for the birds I where hearing. Merlin ID, on the other hand, made me reflect more on acquiring knowledge about the bird, such as species, calls and migration patterns, and trying to find it in the visual environment once I had identified the sound. In this case, the knowledge that I know the bird is here right now made me look more for it. Both applications could benefit greatly from being used in tandem or being integrated, something I will discuss further in chapter 7.1. Another key difference is how the apps support citizen science tasks and integration. Originally, Urban Echoes was intended to be a citizen science project. Still, as the project drifted, this was changed into a project with a higher focus on the aesthetic experiences. Meanwhile, Merlin ID is a great tool for hobbyists to validate their findings before submitting findings to citizen science projects.

2.2.1 Citizen science and bird observation

Citizen science allows regular people to contribute to scientific projects by collecting data on a scale that scientists would otherwise not be able to do[12]. Adopting citizen science within the bird observation community has increased researchers' understanding of bird migration patterns and populations. Since eBirds operate globally, it allows for wider information than regional observation sites like the Danish DOFbasen[17]. Inaturalist[59] is another citizen science application with an accompanying website which focuses on the documentation of wildlife, plants and fungi and thus encompasses a much larger field than eBirds. To keep people engaged, Inaturalist encourages sharing and discussion of nature on the website and application. It lends to a more communal experience than the more individual-focused eBird and Merlin ID. Looking at the real-time map of eBirds, we can see that the global north is overrepresented compared to the global south, indicating a disparity in the amount of data gathered, as seen in Figure 2.1, The light grey dots indicate a recording.

Figure 2.1: A screenshot of the eBird live map, the 31-03-2025[58]

This highlights one potential issue of digitalised citizen science, namely the wealth inequality between the global South and North, which creates a disparity in the number of observations.

An alternative form of citizen science is exemplified by regional bioblitz events, during which scientists and citizens participate in intensive field surveys to document local biodiversity systematically. A research project aimed at creating an app for finding a rare cicada was tested during a bioblitz. The researchers found a strong reluctance toward digital applications among professional and amateur naturalists who participate in bioblitz[35], who preferred more traditional tools such as pen and paper. When asked why they preferred traditional tools over digital, participants mentioned how technology removed them from the nature around them[35].

As mentioned, Urban Echoes was originally intended to have citizen science aspects, but considering the already excellent tools for birding-related citizen science, adding another tool could potentially harm these, as it could split user bases between the two applications. Another concern is that the users contributing to the projects are often deeply engaged in their local wildlife and nature. Thus, they might find the use of technology to diminish these experiences. Lastly, the amount of work that goes into creating a trustworthy database was not considered a priority for this project. Currently, Urban Echoes uses sightings from eBirds to populate the map, along with observations made by users of Urban Echoes. There are very few eBirds observations in Aarhus, but the few that are of rarer birds might excite those who hear them.

2.3 DIGITAL STORYTELLING

With Urban Echoes, I intend for the users to tell a collaborative story with the local environment and birds as actors. Digital storytelling is a great way to provide context and meaning to an aesthetic experience. Audio walks are one such instance where the geographical location provides the context for the audio content, forming a coherent experience. Using the geographic position for context is not unique to audio walks, and examples of different digital storytelling, such as videos and other multimedia content, exist. One such example can be found in *Seven Stories*[38], a research project by Nisi et al., where visual location-specific narratives in the form of videos were used to tell local stories on Madera. They found that the user had problems with orientation and GPS fidelity, relying solely on them to play the videos. This encouraged them to use a marker-based solution for videos instead[38]. When the GPS did not function as expected, the user spent much time looking at the phone instead of the environment. Urban Echoes fully relies on the GPS but does not need the same fidelity as Seven Stories.

It's common for audio walks to have a map showing the user's position and points of interest. As seen in the work of Nisi et al., this can lead to frustrating experiences if the GPS fidelity is low, but a high one is required for the design. An audio guide by Pedersen et al. using only auditory media found that removing the map and guiding the user using sound decreased screen usage and increased both observational gain and engagement[41]. I have decided to have no visual elements for Urban Echoes, keeping the user's attention on their visual surroundings. The audio content, while location and context specific, is intended to be found by walking around without any visual aid from the application. Areas without sound will also provide the user with a story, as the area has few birds or recordings. In the latter case, users are encouraged to walk without headphones or only one to hear their present soundscape and potentially contribute with new observations should they hear a bird. The story told becomes the story of which birds were seen where and when, or conversely, the story of what places they avoid, or that people do not use the app. An imagined scenario could, for example, see a silent forest, as users might have closed the application and instead listen to the soundscape around them. Meanwhile, a winter walk through the same forest could have the user listening to all the birds recorded last spring while the real environment is silent.

The interplay between visual and auditory, and how one affects the other. Looked at how landscapes changed the perception of the sound-

scape[31]. They found that spatial arrangement matters more than individual elements in terms of the perception of the soundscape.

2.4 AUDIO BASED INTERFACES

2.5 DESIGN AS RESEARCH AND DRIFTING

3 DESIGN ANALYSIS

In this section, I will analyse the design and describe the design decisions along the way. The sections appear in chronological order and thus reflect how the project developed throughout the writing of the thesis.

3.1 A DIFFERENT DIRECTION: CREATING A USER-MAINTAINED BIRD DATASET FOR VISUAL RECOGNITION

The original vision for the application differed significantly from its final form, notably in the design of the observation mechanic, while the core listening experience remained relatively consistent. Initially, the observation system was envisioned to function through AI-based recognition of birds using both sound and images, similar to MerlinID. The app aimed to distinguish itself through its citizen science component, incorporating elements inspired by iNaturalist [59], where users actively contribute to datasets for visual recognition.

In this model, users would take a photo of a bird—or a group of birds—and then annotate the image by drawing a bounding box around each subject and labelling them, as illustrated in Figure 3.1. Upon uploading the annotated image, the data would be stored as part of a growing object detection dataset. As a form of positive reinforcement, the user would also receive an "observation" entry as a reward for their contribution that could be experienced by the listener. I wanted to create a user-maintained dataset because, like in other citizen science projects, this approach enables scalability far beyond what traditional methods allow. Moreover, there is a notable lack of high-quality avian datasets for object detection tasks [56]. Visual recognition highly depends on the data used to train the model, with the adage "garbage in, garbage out" illustrating this point. A large, high-quality, and labelled dataset could greatly contribute to the avian object detection field.

As the project progressed, I experienced a larger disconnect between the observation and listening experience, and ultimately decided that the project should focus on one or the other. I decided to focus on the aesthetic experience and nature appreciation over citizen science and

Figure 3.1: Annotated image from the system description document for the initial idea

AI aspects of the project, leading to scrapping these parts for a simpler observation approach.

3.2 AUDIO DESIGN

As an audio-based application, a quality audio design is essential for the user experience of Urban Echoes. When starting the project, I considered the audio quality in terms of sampling rate and bitrate. After discussing my project with audio experts in sound study and sound aesthetics, this perspective changed. They pointed me in the direction of soundscapes[47] and acoustic ecology[47, 57]. Soundscapes and acoustic ecology hint at the interplay between sounds and how they affect each other to form a complete impression, much akin to how each instrument in the orchestra forms the coherent listening experience. In a forest, for example, we might hear the wind blowing, swaying the trees, the small stream trickling nearby, some small animal in the bushes, and the sound of our boots on the ground. All these sounds are heard and perceived together. Transposing a sound file with all its associated acoustic ecology to be played at another place and time we might create a dissonance between the playback soundscape and the limited visual environment. This can happen when our visual environment has a vastly different acoustic ecology, and this ecology is known to us more than the sound playback we are hearing. Imagine walking down a street in a European city while a recording of the Common Swift is playing. The sound of the bird might be familiar, as the common swift can be found in many European cities. However, if the soundscape of the sound files is that of the sub-Saharan rainforest with all its associated sounds, it is clear that the soundscape does not belong to the street you are walking on. As I seek to give users

an affiliation with their local nature through sound, the soundscape should fit with what we expect.

3.2.1 The Challenge of Aligning the Acoustic Ecology with the Visual Environments

To ensure alignment between the visual environment and the soundscape. I wanted to isolate bird vocalisations from their ecological background so that users could either immerse themselves purely in these vocalisations or, by enabling transparency mode on their devices, allow the real-world soundscape to blend naturally with the playback audio.

I initially experimented with a simple noise gate to isolate bird vocalisations and filter out most of the ecological background to achieve. Unfortunately, this approach was ineffective due to the dynamic range of the bird calls, which led to parts of the calls being inadvertently cut off. I then considered limiting playback to sound files from the user's country. However, this would significantly reduce the app's applicability in countries with limited available recordings.

I also explored the use of AI-generated audio via the ElevenLabs text-to-sound effects (SFX) feature [20]. Despite being promising in theory, the resulting bird calls lacked realism and were downright awful at worst. Ultimately, I reverted to using high-quality, category A recordings from the Xeno-Canto database, selected at random. Although not perfect, this method provided the most consistent and natural listening experience. A final option considered was to train an AI model capable of recognising and removing background noise. Although potentially effective, this was deemed too time-consuming for the current project. Nevertheless, implementing such a filter remains a compelling direction for future work.

3.2.2 Evaluating the sonic aesthetics

To evaluate the Sonic aesthetics, I used a framework by Cunningham et al.[16]. In their view, sonic aesthetics can be evaluated through the following criteria *intensity, pitch, timing, spatial, fidelity, context, originality, and expectation*[16]. The definitions of these dimensions can be seen in the figure 3.2 taken from their paper *Towards a Framework of Aesthetics in Sonic Interaction*[16]
Looking at this framework, Urban Echoes fulfils most of the criteria for highly aesthetic sonic design with a few exceptions. By choosing only sound files with the highest quality rating from Xeno-Canto, and by the qualities of bird songs and calls as aesthetic sounds, we can,

Dimension	Description
Intensity	The perception of loudness of the sound.
Pitch	The fundamental frequency, or musical note, dominant in the sound.
Timbre	The character of the sound and complexity of its frequency spectrum.
Spatial	The position of the sound relative to the listener's position.

(a) Dimensions: Intensity, Pitch, Timbre, Spatial

Fidelity	The production quality of the sound. Considers the clarity of the sound and its intention, including the presence of noise and other artefacts and the discernibility between signal and noise.
Context	The relationship between the sound, interaction task, and any other sensory stimulations (e.g., touch, visual, olfaction). Context may not always be known to the designer (such as for mobile applications).
Originality and Expectation	The level of predictability and familiarity in terms of the overall context of use, as well as in general experience.

(b) Dimensions: Fidelity, context, Originality and Expectation

Figure 3.2: The definitions of the dimensions from the paper *Towards a Framework of Aesthetics in Sonic Interaction* [16]

for most sounds, expect a good timbre and pitch. The intensity is mostly good, but varying loudness of sound files can lead to lower aesthetics, as some files might be too loud while others are too low. Initially used to design and argue for the audio, I decided to include these as evaluation parameters for my test.

I believe both spatially, context and Originality and expectation to be good, am unsure about Fidelity. Could be used as evaluation parameters, if so, should I include the results here?

For my vertical slice of the prototype discussed in Chapter 4.3, I found specific sound files with as little background noise as possible to transpose them. However, for the final prototype,

3.3 DESIGNING THE LISTENER ROLE

The application's listening experience lets users explore the auditory map. The map is explored by walking or otherwise moving, location tracking will work together with audio playback features to play the observations of a given location. During the design of the listening experience, I reflected on what maps can tell us and how we use them.

3.3.1 Mapping the world

When creating a map, we must understand the information we are mapping and how the readers will interpret the map. Questions such as, What are the intentions behind the creation of the map, and What kind of world does the map show? Take, for example, the most classic map projection, the Mercator map projection, where angles are preserved at the cost of inflating the areas of landmasses further towards the poles while minimising the ones closer to the equator. The angel-preserving properties make it useful for nautical navigation, but its ubiquitous application has given rise to a common critique against the projection in how Europe and North America are inflated in a colonialist manner[34]. Bearing these questions in mind helped me focus on what kind of world Urban Echoes mapping would reflect. Since it's meant to be used predominantly within an urban setting, it does not reflect where the most birds can be found. Instead, it will show where users choose to engage with the app. Cities, in this case, could be overrepresented compared to rural areas, with fewer users but possibly larger bird populations. Another aspect to consider in this regard is how the app's functionality relies on a stable connection. Therefore, areas without an Internet connection cannot be experienced or observations made using the application. This can be misleading if I do not clarify that this is not a direct mapping of the local birds.

This view of the interaction as something not just between the application and the user, but the wider circumstances surrounding the design, is also found in the literature, such as with Alex Taylor's "after interaction"[52], where the use of rental bikes paints a unique map of use.

Similarly to how the Mercator projection is useful for nautical navigation, I wanted to investigate the activity Urban Echoes bird mapping would create. As a listing experience, the map is intended to get people to pay attention to their visual environment while immersing themselves in the sonic. Experience the bird where it was seen, even if it is no longer physically there. In places without observations, silence can serve as an invitation to find birds, so observations can be created.

3.3.2 Mapping sounds

Used as a sound of maps[53] Either remove or expand

The sound enhanced the traditional map viewing experience[28]. They augmented a digital map of a hiking area by mapping sound files to areas of the map so that when users move their cursor above a certain area, like the forest would hear the associated sound. This helped

people with impaired vision gain something from the map viewing experience and improved the general perception of use.

How big of a discovery radius should an observation have? I had to decide how large the radius for an observation should be; in the end, I decided to go for a 50-meter radius. After testing out some different options, I found that 50 meters gave the best experience in regards to how long you hear the sounds.

3.4 ADDING TIME TEMPORAL ASPECTS SEASONS

I wanted to include temporal aspects in both time of day and season. In learning about birds, I found that these aspects could also be taught to users. There is a connection between breeding cycles and the amount of singing[50]. In Europe, birds often breed around the time when food is the most abundant for their offspring[29]. In northern Europe, birds breed most often in late spring or early autumn[29]. Beyond this, some birds migrate depending on the seasons, which makes the frequency and type of bird vocalisations depend on the season. I wanted the application to showcase the difference in seasons and allow a user to, for example, hear how a place sounds during spring in early December. Therefore, observations are loaded based on seasons. Users can choose which seasons they want to listen to at a given moment, with their current season being the default. I implemented this, but it did not affect the prototype testing. These seasonal changes would require the prototype to be used longer before becoming noticeable. I also wanted to include the effects on time of day, as birds are most active at and around sunrise[45], with activity the lowest during the nights and a local minimum at midday. However, I choose not to implement this for the prototype as I want as high a chance of the evaluators encountering each other's observations. Thus, limiting evaluations to certain times of day would make the chances of this less. In an idealised version, however, the time of day would be reflected in the observations, with the user potentially being able to choose when to listen to the particular area.

3.5 CREATING A FULLY AUDITORY APPLICATION

The application was interested in

4 DESIGN AND EVALUATION OF SYSTEM ARCHITECTURE

In this section of the Thesis, I review the prototype's system architecture and changes between major iterations. Finally, I evaluate the iterations and how they affected further development.

4.1 GENERAL SYSTEM ARCHITECTURE

Some general description and a UML diagram, maybe

4.2 FLUTTER AND DART

Most of the application is written in Dart, a language created for mobile and web development with C-like syntax. The choice to use Dart came from a desire to use Flutter[22], a multiplatform open-source framework developed and maintained by Google. I wanted to use a multiplatform framework so that the application could run on Android, iOS, and the web with a single codebase, even if I later decided not to develop a web application. There was no particular reason for choosing Flutter over other multiplatform frameworks like React Native, aside from a personal desire to learn a new language and framework.

Android was chosen as the main development platform, so the app was tested and implemented primarily on Android devices. This decision was based on the more lenient deployment policies compared to iOS and the fact that I own multiple Android devices but no iOS devices.

4.3 VERTICAL SLICE OF THE PROTOTYPE

I still believe this section would benefit from discussing design as research as part of the introduction. Alternatively, maybe I could discuss design as research elsewhere. probably makes most sense in related work or introdcution

To experience how the prototype would feel, I tested a straightforward small-scale implementation version capable of playing the observa-

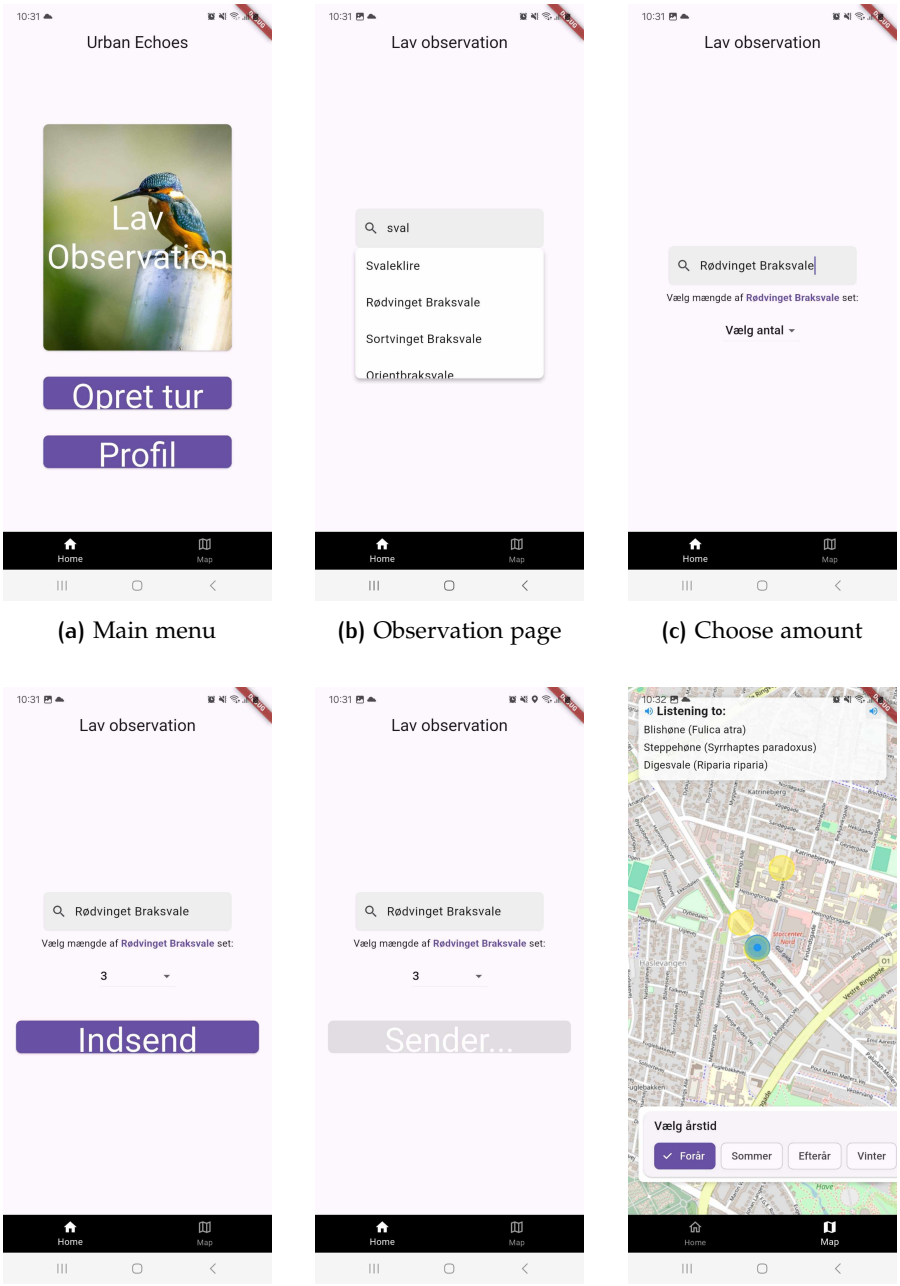
tions to create a listener experience.

This presents itself as a vertical slice of the final application capable of all the features related to the audio walk part, but lacking features such as observation recording, logging on and sending, and tracking the trip.

To find out what kind of birds I would use for the vertical slice prototype, I went on a walk in the area and listened to the sounds of the birds. I found that the neighbourhood of the villa, where there are many large gardens with trees, had significantly more birds. Observing the birds, I also found a lot more than when going for a regular walk, highlighting the potential of the observer role as it encourages greater sonic awareness to identify the birds. I found that the birds at long periods did not produce any sounds, which I initially wanted to emulate for the prototype, before I decided on a more exaggerated approach. Additionally, I thought about the radius, and I think somewhere between 30 and 50 meters seems fine without having tested it further. These sounds drown out the birds near the road or during construction work.
The birds I encountered during my short walk were the following:

- Eurasian Blackbird (Turdus merula)
- European Herring Gull (Larus argentatus)
- Common Wood Pigeon (Columba palumbus)
- Eurasian Blue Tit (Cyanistes caeruleus)
- Rook (Corvus frugilegus)
- Common Gull (Larus canus)
- Eurasian Tree Sparrow (Passer montanus)
- European Greenfinch (Chloris chloris)

These were inserted as points into the database with the relevant coordinates and times. A visual interface was provided for debugging purposes to allow me to identify problems with the prototype more easily. Figure 4.1 shows a view of how the visuals looked. Write a description of the visual
I tested the vertical slice prototype with my advisor and later with a fellow IT product development master's student. A bug related to the GPS functionality somewhat hindered the prototype's evaluation, with the GPS updating infrequently causing certain observations to play for an extended time and others to play at all. However, both tests allowed me to reflect on the successes and what needed to be improved. One finding related to the exaggerated nature of the calls in the application is that when compared to real-life sightings, audio recordings would be both louder and of higher intensity than the actual calls and songs. The GPS problems came from the screen timeout that the phone would enter after a short use time. I originally tried to solve this by allowing



(a) Main menu **(b)** Observation page **(c)** Choose amount

(d) Send in the recording **(e)** Sending the recoding **(f)** The listener view

Figure 4.1: The visuals for the vertical slice

audio and GPS to run in the background. But while I successfully got the audio player to work as a background service, the GPS kept giving me trouble until I decided to have the app prevent the phone from going into Screen timeout as long as the app is in the foreground. Although preventing screen timeout consumes a lot of battery life in my already battery-intensive application, I decided that this was a sacrifice needed to move on to other areas of development.

4.4 VOICE RECOGNITION AND AUDIO

Initially, I implemented a text-to-speech solution using the Flutter package called *flutter_tts*[18]. This solution used prewritten text pieces, which would be read out when certain conditions were met. I preferred this solution over a pre-recorded audio solution, as adding more birds and languages would be quicker. The app is built for a Danish audience; therefore, the audio language is Danish. Unfortunately, the pronunciation of Danish words with the text-to-speech model sounded almost intelligible. This made me change over to a pre-recorded solution. I recorded the bird names individually to allow them to be inserted into multiple sentences. This made it possible to create dynamic sentences such as (translated to English, though the app uses Danish): "Did you see a" followed by a recording of the bird name, resulting in complete phrases like: "Have you seen a Blackbird?" albeit with a small delay. This also meant the application doesn't have voicelines for every bird found in Denmark, as it was very time-intensive to do the recordings. However, 53 of the most common birds out of the roughly 300 birds in Denmark have been recorded. To test the new functionalities, I made a new debug interface which can be seen in Figure 4.2. These visuals ended up being the basics for the prototypes final look as seen in 1.1.

4.5 BACKEND

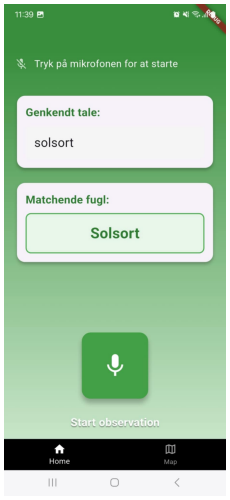
The backend is built using Python and FastAPI and hosted on Microsoft Azure. Microsoft Azure was chosen over other cloud hosting services for several reasons. The service scales well and offers all the features needed for my backend, such as web hosting, PostgreSQL database support and blob storage. Furthermore, students get 100 dollars worth of credit through the GitHub Student package, which means I could use the hosting without paying out of pocket. The choice to run a FastAPI backend was mainly based on my familiarity with it doing previous projects.

4.5.1 Database

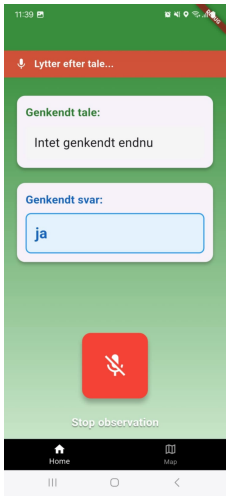
The PostgreSQL database has two tables: `bird_observation_table`, seen in Table 4.1, and `birds`, seen in Table 4.2.
Furthermore, an Azure storage blob holds sound files for all birds found in Denmark. Each bird has 20 associated sound files downloaded from Xeno-Canto. The blob directories are associated with the Database. When an observation is active, the application will stream and play a sound file associated with the link on the blob. For a less data-intensive solution, the app could include all the sound files for

(a) Listening is triggered by clicking the large green button.

(b) The app is listening and has matched the spoken word with "solsort" (Blackbird).



(c) The recognised speech and matched word are displayed.



(d) The user says "ja" (yes) to confirm the observation, prompting the system to save it in the database.

Figure 4.2: Screenshots illustrating the speech recognition debug visuals in the application.

Field	Type	Description
id	INT	Unique identifier for each observation
bird_name	VARCHAR	Common name of the bird
scientific_name	VARCHAR	Scientific name (e.g., *Corvus corax*)
sound_directory	TEXT	Path to the associated sound file
latitude	numeric	Latitude of the observation point
longitude	numeric	Longitude of the observation point
observation_date	DATE	Date of the observation
observation_time	TIME	Time of the observation
observer_id	INT	Reference to the observer
quantity	INT	Number of birds observed
is_test_data	BOOLEAN	Flag to indicate test or real data
test_batch_id	VARCHAR	ID of the test batch (if applicable)
source_id	TEXT	ID from eBirds on eBird observations

Table 4.1: Schema for the `bird_observations` table

Field	Type	Description
id	INT	Unique identifier for the bird
common_name	VARCHAR	Common name of the bird
scientific_name	VARCHAR	Scientific name (e.g., *Turdus merula*)
danish_name	VARCHAR	Bird's name in Danish
region	VARCHAR	Region where the observation was made
lastobserved	TIMESTAMP WITHOUT TIME ZONE	Date the bird was last observed
is_common	BOOLEAN	Whether the bird is commonly found

Table 4.2: Schema for the `birds` table

regional birds or have them available for download. The app will still need an internet connection as it will need to access points, but these could also be cached, further decreasing the data intensity.

4.6 SPEECH RECOGNITION, TEXT TO SPEECH AND RECORDINGS

When the user clicks the microphone icon at the bottom of the screen, they will go into listing mode as seen in figure 4.2b. While in listing

mode, the application will take the user's speech as input and use the Flutter package *speech_to_text* v7.0.0 [14] to turn it into text. The text will then be matched against a species list, and if a match is found, the system will proceed by prompting the user with the confirmation card as seen in figure 1.1b. Special words such as the seasons and ja(yes), nej(no), and other confirmation words also get's checked for matches. Depending on the match, the application can enter different stages, such as the confirmation stage, after the user says ja(yes) to having observed a bird. The list of birds is taken from the species list from netfulg.dk[36] ensures that every species found in Denmark is included in the application. Since the app is intended for a Danish audience, both the speech recognition matches and audio playback are in Danish.

4.6.1 Text to speech

I implemented a text-to-speech functionality in the application for announcing bird names and playing a few voice lines, such as "Have you observed a (bird name)"[1] and "Observation for (bird name) has been created."[2]. The text-to-speech was implemented using a library called *flutter_tts* v4.2.2 [18]. After testing the library's Danish capabilities, I found that it struggled with pronunciation to the point of it becoming almost unintelligible. I therefore decided to go for an adioplayback solution instead, where I would record voicelines.

4.7 AUDITORY CONSIDERATIONS AND HOW IT TRANSLATES TO CODE

1 Actual sentence in application said in Danish: *Har du observeret en (fuglens navn)*
2 Actual sentence in application said in Danish: *Observation for (fuglens navn) er oprettet*

5 IMPLEMENTATION

In this section, I will go more in-depth with the implementation of the different components of the final prototype. The application's general architecture follows the Provider pattern to establish a hierarchical dependency injection system. Listing 5.1 illustrates how various provider types manage service dependencies, ranging from basic stateless services to complex coordinating services that orchestrate multiple components. The full setup is provided in Appendix B. A provider pattern was chosen due to the separation of logic through dependency injection, making the system more manageable, modular and understandable.

```
MultiProvider(
  providers: [
    // Core state managers
    ChangeNotifierProvider<PageStateManager>(create: (context) =>
        PageStateManager()),

    // Basic services
    Provider<DatabaseService>(create: (context) =>
        DatabaseService(), lazy: false),

    // Services with dependencies
    ListenableProvider<ObservationUploader>(
      create: (context) => ObservationUploader(
        databaseService: Provider.of<DatabaseService>(context,
            listen: false),
        // Other dependencies...
      ),
    ),

    // Complex coordinating service with multiple dependencies
    //Speech coordinator
    ChangeNotifierProxyProvider4<...>(...),
  ],
  child: MaterialApp(...),
)
```

Listing 5.1: MultiProvider setup in Flutter

5.1 AUDITORY MAP(LOCATION SERVICE)

The auditory map is created and managed by a class called location_service with the following dependencies

- bird_sound_player
- azure_storage_service
- background_audio_service
- season_service
- location_manager

Before delving into how this is done, I wish to describe the dependencies so that a greater understanding of the system is gained.

5.1.1 bird_sound_player

This module initialises a pool of five audio players(the exact number is set in the service_config file), manages which players are available, and handles the playback of bird sounds using the available players. The audio players used are from the Flutter library audioplayers version 6.2.0[37].

5.1.2 background_audio_service

This class ensures that sounds keep playing even when the application is minimised or the screen is turned off. Unfortunately, sounds will not change while the app runs in the background, as I were not able to get location tracking working as a background service.

5.1.3 season_service

A simple class for tracking the seasons. This means keeping track of the current seasons as well as finding seasons for other observations based on their dates. I decided that the application should allow users to change seasons and then hear only observations made in that season. The class is not used in the prototype as I removed the feature to change between seasons, as there were so few bird observations and since the participants in the evaluation would all be using it during the same season.

5.1.4 azure_storage_service

This class followed a singleton pattern and has two responsibilities. Firstly, it should initialise the connection to the Azure storage blob holding all the sound files.

```
Future initialize() async {
  try {
    if (_initialized && _storage != null) return true;

    _storageAccountName = dotenv.env['AZURE_STORAGE_ACCOUNT_NAME
        '] ?? '';
    final connectionString =
        dotenv.env['AZURE_STORAGE_CONNECTION_STRING'] ?? '';

    if (_storageAccountName!.isEmpty || connectionString.isEmpty)
         {
      debugPrint('Azure Storage credentials are missing');
      return false;
    }

    _storage = AzureStorage.parse(connectionString);

    _initialized = true;
    debugPrint('Azure Storage Service initialized successfully');
    return true;
  } catch (e) {
    debugPrint('Error initializing Azure Storage Service: $e');
    _initialized = false;
    return false;
  }
}
```

Listing 5.2: Azure Storage Service Initialization Method

Its second responsibility is to list all the files in a certain folder of the blob storage. This is used as the bird sound plays a random sound file from a certain folder when playing a bird sound.

5.1.5 location_manager

The `LocationManager` class is responsible for tracking the user's location. It uses the `LocationRepositoryInterface` as an abstraction layer over the Flutter package `geolocator`[4].

```dart
void _handlePositionUpdate(Position position) {
  _currentPosition = position;

  // Skip processing if not enough time has elapsed
  final now = DateTime.now();
  if (now.difference(_lastPositionUpdate).inMilliseconds < 500) {
    return;
  }
  _lastPositionUpdate = now;

  if (_config.debugMode) {
    debugPrint('[LocationManager] Position update: ${position.
        latitude}, ${position.longitude}');
  }

  // Skip if position hasn't changed significantly
  if (_lastProcessedPosition != null) {
    final distance = _locationRepository.distanceBetween(
      _lastProcessedPosition!.latitude,
      _lastProcessedPosition!.longitude,
      position.latitude,
      position.longitude
    );

    if (distance < _config.distanceFilter / 2) {
      if (_config.debugMode) {
        debugPrint('[LocationManager] Skipping position update (
            moved only ${distance.toStringAsFixed(1)}m)');
      }
      return;
    }
  }

  _lastProcessedPosition = position;

  // Call the callback
  if (onPositionUpdate != null) {
    onPositionUpdate!(position);
  }
}
```

Listing 5.3: Dart Function: _handlePositionUpdate

5.1.6 How it all works together(Need more in depth rewrite now that I have explained all the dependencies)

The map works by making a call to the backend that retrieves all the observations of that season. As each point has an associated geographical position. During the debugging process, a visual map as seen in figure 4.1f showed observations and user location along with what birds the user was hearing. The map was created using flutter_map[13]. I choose to remove the visuals of the map so users don't just look at the phone instead of the environment. Similar reading goes for the widget that showed what kind of birds the user where hearing.
New sightings are fetched at startup
use the eBird[51] to populate the database

5.2 OBSERVATION UPLOADER AND DATABASE CONNECTIONS

5.3 RECORDINGPLAYERSERVICE

5.4 SPEECH COORDINATOR SERVICE

5.5 BACKEND

A FastAPI backend written in Python and hosted on a student subscription to Azure provides an endpoint for fetching observations. There are other endpoints, but these are not used in the version of the prototype that was evaluated. The backend can be found in Appendix B.2. As part of the backend, a Postgres database holds a table for observations. Additionally, I run an Azure storage blob for the sound files. All sound files have been collected from xeno-carto[23] and are organised in folders based on the scientific name of the bird. These folders are associated with the observations in the database, allowing me to fetch random sound files from specific folders based on the bird observed in a given observation. I use scientific names for the back-end as they are language-independent, making it more scalable.

6 EVALUATION

In this section, I describe the evaluation starting with the methodology. Afterwards, I present the results of the evaluation, followed by a discussion section.

6.1 METHODOLOGY

The prototype was evaluated over a week, with most participants starting the evaluation on the 12th of May and ending on the 19th of May. Three evaluations ran from the 14th to the 21st and were scheduled later, as not all participants had time to start the evaluation on the 12th. During the week, each evaluator were asked to use the application at least once every day over a week. After each use, they were also asked to write down their thoughts with an emphasis on writing, preferably 5 lines or more each day. Lastly, they were asked to write some final reflections about how it was using the app over the week. This study is inspired by diary studies[21] but does not feature prompts. I decided not to include any prompts to allow the evaluators to bring their unique expertise into play, like, for example, their technical backgrounds. However, by not including prompts, I might have found fewer similarities in answers than if I had done otherwise. The resulting diaries, both original and translated, can be read in the Appendix A.1.

Each evaluation started with the participant downloading the application. Afterwards, the evaluator and participant went on a short walk where participants learned the to use the app and were able to ask questions. This walk was also meant as a way to be able to, discuss participants' initial thoughts. After each walk, a small semi-structured interview[1] of approximately 15-30 minutes was conducted. All interview transcripts are in the Appendix A. Here, the evaluator was able to grasp the participants' initial experiences with the application and how much of an existing interest they had in the local bird wildlife. Furthermore, the interview serves as a starting point for reflection for the participants.

6.1.1 Participants

The study involved ten participants, aged 21 to 61, comprising four women and six men. The majority were students, with eight participants currently enrolled in study programs. Among these, three were pursuing a Master's degree in IT Product Development. The remaining students were enrolled in a Bachelor's program in Medical Chemistry, a Master's program in Economics, a Bachelor's program in Chemistry, and a Professional Bachelor's program in Diaconia and Social Education. The other three participants were employed: one as a Ph.D. candidate in the Department of Chemistry at Aarhus University, one as a system developer in the Department of Physics, and one as a pedagogue at a school. Of the participants, two were already interested in birds, with most of the participants waiting to learn more about their environment.

In addition, an eleventh individual participated in a limited capacity due to their Android device running an outdated operating system, which restricted them to using only the listing activity of the experience. This participant is currently pursuing a Bachelor's degree in Biology.

6.2 RESULTS(THIS SECTION IS SUBJECT TO CHANGE AS ALL THE EVALUATIONS ARE NOT FINISHED YET)

The results have been split into three sections: the interview section, which consist of my reflections on the interview answers, a section showcasing the results of the diary study, and lastly a section presenting the results written as part of participants final reflections.

6.2.1 Interview results

Participants 1 and 5 were not interviewed, and as such, the interview section consists of only 9 participants.

8 Participants had an overall initial positive response to the application, and all 9 participants noticed more birds than when going for a regular walk. This difference could be due to how they were actively looking for the birds to record them, as well as looking for the birds they were hearing through the app. All 9 participants used the Merlin ID app to help them identify the birds. However, the 2 participants who had a pre-existing interest in birds, for the most part, did not use Merlin ID but used it when unsure about what kind of bird they had seen

or heard. Multiple participants asked about combining the functionality of the apps so did not have to change between apps all the time. Something which became a reappearing trend in the dairy study result

Several participants reported going on regular walks, often seeking out natural environments during these outings. They described the walks as relaxing and helpful for clearing the mind. For some, the soundscape played an important role in this experience—natural sounds such as birdsong and the wind in the trees contributed to a sense of calm. Others, however, preferred to listen to music while walking. In contrast, many participants found the urban soundscape to be noisy and overstimulating, leading to discomfort when walking near roads or in areas with significant industrial, mechanical, or human-made noise. One participant, however, occasionally appreciated industrial areas—particularly when they had a certain aesthetic cohesion. For example, he described enjoying the atmosphere of the harbour, where even a boat's horn—typically considered an unpleasant sound—could become a positive part of the experience. He clarified, though, that this appreciation did not extend to traffic or road noise.

One participant found it hard to identify which bird sounds were real and which were from the app, forcing him to rely on visually seeing and entering the birds rather than identifying them through sound. This sentiment was shared by another participant who said the experience was somewhat confusing and overwhelming, as they had to relate to both the digital birds and real ones. They did, however, find it intriguing and fun to have a digital element to the bird-watching experience.

One participant was concerned about primarily hearing certain types of birds, namely pigeons and seagulls, as they are quite common in Aarhus. The participant dislikes these birds. When asked why the participant thought this way about the animals, he cited seagulls as food thieves and aggressive, calling their calls unpleasant to hear. As for pigeons, when he was asked why he disliked them, he somewhat changed his stance and mentioned how he thinks of them as pests, but not as bad as the seagulls. He finds their song/call to be pleasant, but still does not find it as pretty as, for example, the blackbird's song.

Some general concerns with technical issues were shared amongst many participants, with issues such as the playback voice being low compared to the bird sounds, and the record button sometimes being a bit unresponsive. One participant in particular had a lot of issues with the observation mechanism, finding the speech interface slow and bothersome and wished that they could just use the search system included as a last resort as the only method for uploading. This

participant also faced more difficulties with uploading, with uploads often failing throughout the entire week, something that would rarely happen to other participants but would happen almost every third time for this participant. They also got so frustrated with the speech recognition that they made a false upload during the initial walk after the system heard the bird wrong.

A participant commented on how they could see the application being used as a pedagogical tool to engage people with special needs, such as individuals with Down syndrome, in bird wildlife. They attributed the app's simple interface and task-based/playing interaction of finding and recording birds as observations.

As I walked with them, I also made some observations myself. Most participants walked with the phone out and looked at it when making observations. When not, they were actively looking around the environment, often stopping to identify or look for a bird.

6.2.2 Results from the Dairy study

Questions for Peter: Also, is it fine to have so many quotes in this section, or should I do less? Right now, I am considering writing a little before each and then referring to them in the discussion.

A common finding amongst participants was an increased awareness of both sonic and visual elements in their environment. The increased awareness led the participants to notice more birds than usual. It also became apparent that participants sought to connect their auditory and visual environments. A general positive attitude towards this increased awareness was also present in these responses, of which a few can be seen below.

> "It was a 40-minute walk through the park area (in Skejbyparken). It was quite a lot of fun to figure out which bird it was, and guess before checking. I also felt like I more actively looked for them when hearing them, linking sound to visual." - Participant 7
>
> "This time, I chose to go for a walk with the app late in the evening. I only walked in the Urban areas, and it was so late, so I didn't see Any birds around me, but I just listened to the app.
> I think using the app, I looked more around me as I was trying to see if I could see any of the birds I was hearing through the app.

> It was pleasant to look more around me while not being as disturbed by the noises of cars with the headphones in." - Participant 4
>
> "Within about 1 km of Børglum Kollegiet: In the immediate area of Børglum (< 1 km), I was able to hear birdsong through the app. I liked how the resulting attention towards possible bird sightings made me more focused, and gave me a "mental pause" in that way, more so than music would usually do." - Participant 8
>
> "It feels nice to hear the ambient sound in the headphone(s) almost perfectly match what I could hear IRL. I noticed myself taking longer than usual just taking in all the sounds of nature, the birds, trees rustling or creaking, a law mover(lawnmower) in the distance. All together I feel refreshed." - Participant 7

Participant 4 also noted how it helped her pay attention to the sounds of nature, something she normally struggles with in cities due to the noisy soundscape.

> "I found the app made a notable difference in the more urban environments, where it was very easy to spot the birds but hard to hear them. When the birds were flying in the sky and between buildings, they stood out easily, but due to noise pollution from cars and people, it was hard to hear them. Especially as an autistic individual, it is very hard not to get overstimulated by the noises in the city and not even enjoy the sounds of nature at all. Using the headphones with the app, I was able to synchronise the sound of the birds as I saw them." - Participant 4

Participant 9 became possibly surprised by the amount of birdlife in the city and the diversity in it, commenting on it both when using it in Aarhus and their home city of Søften, a suburb of Aarhus located roughly 15 kilometres northwest of Aarhus. Interestingly, this was something he had never noticed before, but was easily detectable after actively looking for the birds.

> "Wild plants also grow there, which benefits biodiversity. This is also reflected in the birdlife. Although the area is not nearly as large as the Botanical Garden, the same bird species are present. It's nice to see such diversity in different places within Aarhus Municipality." - Participant 9 (translated from Danish)
>
> "First time using the app in Søften, where I live. Through Urban Echoes and Merlin Bird ID, I've discovered that there are many more bird species in my local area than

> I previously thought. It's a pleasant surprise that bird diversity is thriving where I live. I've lived in Søften my whole life, and only now am I getting insight into the birds I didn't know we had." - Participant 9 (translated from Danish)

Using Urban Echoes, participants identified which birds were common and where they reside, but this familiarity also led to fatigue for participant 3.

> "On the way there, I noticed how many of the same kinds of birds there are in Aarhus. It started to get a bit less exciting to enter blackbirds and wood pigeons over and over again. But then it suddenly became exciting when I walked around a corner and heard a great tit and a jackdaw — you get a little rush from finding a new bird." - Participant 3(translated from Danish)

Participant 4 started reflecting on how the knowledge of which birds are common and where they reside could be used for city planers and communities to help birds of a certain area, as seen in the quote and Figure 6.1.

> "I've been noticing that within the more Urban areas, the birds seem to have a large preference for one type of environment, therefore making the type of birds you'll see or hear in certain environments very low in diversity. Today I walked through a small neighbourhood with, honestly, a large amount of birds for an Urban area, but they were 95 per cent just one kind of bird. Ive noticed the same in another neighbourhood with another type of bird. Maybe it says something about their preferred environment, or maybe they choose to flock together in certain areas, like a neighbourhood. Either Way, it made me think that you could possibly set up amenities for the birds that the app plots in certain areas to encourage birds to live within cities" - Participant 4

Several participants commented on being uncomfortable with having to say bird names out loud in public for the interaction, feeling that it drew unwanted attention.

> "It does feel a little awkward to have to relatively loudly say the name in public when alone (if I had someone with me it would feel less awkward to talk out aloud), especially when others are walking by." - Participant 7
>
> "I went for a walk in the botanical garden. It was windy, so the app had trouble recognising my speech, and I became

Figure 6.1: A quick sketch by participant 4 drawn after explaining her idea for bird neighbourhoods

> quite self-conscious when I had to say bird names out loud in public, especially when I had to repeat myself." - Participant 10 (translated from Danish)

The use of the app helped Participant 7 increase their knowledge of birds. They found it enjoyable to begin recognising birds on their own and mentioned that they no longer needed Merlin ID to identify every bird. This sentiment was also shared by Participant 3 during the final reflections.

> "It is kind of fun to slowly recognise the birds without needing an extra app. It feels like I am learning something about my local fauna." - participant 7

Participants 5 and 11 were in disagreement on whether or not having to say the bird's real name instead of its common name was annoying or beneficial. Participant 5's reflection also showcases a fault in the suggestion section, with the suggestion algorithm giving results that are too far removed from the spoken word.

> "I heard quite a few birds and also saw some. It would be smart if the app could recognise birds' common names and suggest the correct species. For example, the hooded crow is usually just known as a crow. If you say crow, many suggestions appear, but not hooded crow. If a crow showed up, it would be easier to select the right bird." - Participant 5(translated from Danish)

> "Again, learning birds' real names – it's good that you're forced to give a name to a bird in the database." - Participant 11 (translated from Danish)

Participant 7 thought of a share feature where they would be able to share what they are hearing with someone not located at the same geographical position. The share feature was something also originally envisioned as part of the application, but was not implemented for the prototype.

> "Something that came to mind when making new observations and hearing the ones from yesterday. Is that it could be great if I could share the observation, the sound ambience, with someone who isn't in the same locale as me. To share with them how it sounds to be here in nature." - Participant 7

Several participants recorded frustrations with the technical aspects of the application. Issues related to the microphone button, speech recognition and upload difficulties plagued many participants' experience. A critical bug related to the confirmation card not appearing with some birds proved so detrimental that an update fixing this problem was pushed on the second day of the evaluation, and all participants then had to update.

> "Record button does not change back to green but stops using the microphone. With a repeat press (while it is red and inactive) will turn the microphone back on. So it functions as it should but visuals are off. It was windy so it was very hard to log observations. And therefore I think a search option from the start menu would be a huge quality of life feature. Another reason I would love this feature is because I hate making verbal observations in public. As such, if I was a normal user, the lack of this feature would be enough for me to stop using the app entirely. So i think giving the user the option to chose to make the observation though verbal or written means would be an improvement to the app." - Participant 6

> "The app starts listening again once the first observation is registered, which is annoying." - Participant 10 (translated from Danish)

> "The spoken response to accept a bird feels as a rather slow way to acknowledge the logging. Spoken responses would have made sense if you could accept the logging by answering 'Ja' og 'Nej' – but when you have to accept the logging by pressing the screen it would have been much faster with just a text on the screen for this." - Participant 2

It should be noted that the app does allow for users to answer with 'Ja' and 'Nej'; however, this participant had so much trouble with it that they believed it was not a feature.

Participants 3,4, and 5 all used the application during the late evening and nighttime. Their reflections point towards how different times of day affect the user experience.

> "The bird noises from the app didn't match up with the behaviour of the birds in this instance. Although it is obvious that the birds would need to rest, the behaviour seen in real life vs heard through the app was extremely disconnected." - Participant 4
>
> "The app wasn't very useful, as no birds were awake and there were no bird observations in the app for that location yet." - Participant 3
>
> "I went for an evening walk to Netto on Randersvej. It was dark, and I didn't notice any birds. I think they were asleep." - Participant 5

Gamification and tracking of birds found.

> "It was still fun to discuss with the others which birds we could see and hear. I really enjoy discovering new birds—I heard a chiffchaff, a greylag goose, and a bird whose name I've just forgotten, but which I hadn't found in Aarhus before. It gives you the feeling of collecting Pokémon. I think it would be cool to have a Pokédex-style feature where you could see the birds you've 'collected.' Something that makes it a bit more gamified." - Participant 3
>
> I found that I needed the Merlin Bird ID app to figure out which birds were present today. I also found that I missed the option to see which birds I myself had added to the database."

6.2.3 Final relections

> "In general, most of my tours were in nature like areas with lots of bird sounds at this time of the year. In this use of the app the bird sounds - which are working fine - would have made more sense for me if the app displayed the name and a picture of the bird. I know that this is probably not the intention of the app - as the app name says, it may probably make more sense to add bird sounds to more urban areas that I have been walking in." - Participant 2

> "However, just walking with the app made me more aware of the bird sounds - especially the great number of different birds. The process of registration of the birds in itself also helped increase the awareness of the many different birds." - Participant 2
>
> "Apart from that, I can see a lot of potential in location based sounds/oral information. And also the potential in letting users be part of generating the information/sound landscape." - Participant 2

6.3 DISCUSSION (FAR FROM DONE)

The application succeeded in making people pay more attention to both the auditory and visual environment. Several participants experienced excitement in finding, observing and hearing new birds in the local environment.

The temporal aspect is something users need more control over. As one user remarked, the sounds can create a disconnected experience if all the birds around the user are asleep. If users were allowed to choose the time of day they wanted to hear, this could be remedied. Furthermore, the evaluations showed a difficulty observing birds later became more difficult. The importance of the connection between the visual and auditory environments. Too big of a disconnect

Multiple participants requested some way to gain more information about the birds they recorded on the app or the ones they heard

I think there is a discussion about how we simply exist with the birds. We pay them little attention, and they often blend into the background; however, with very little effort, we start to notice the birds, something which the participants found pleasant.

Time of day is an important thing to implement in the final version of the application.

Unsure if it could be used as citizen science, with many more observations than eBirds, as people observe the regular birds.

Look at the earlier framework for aesthetic sound and see if you can use it

7 CONCLUSION

In this section, I

7.1 FUTURE WORK

To use the sightings as a reliable database, it's necessary to have verification of the sightings. A few approaches to this could be the use of cross-referencing the sightings with eBird and other users. A user with a history of making high-credibility observations could have a higher certainty measure when uploading an observation. Another approach for verification could use AI for sound identification or image recognition, but this would also force users to either take a picture or a recording.

Another interesting avenue of work would be integrating the features of Urban Echoes into large birding applications like MerlinID or GoBirds. Integrating UrbanEchoes into MerlinID, for example, would allow observations to be created when the user records a bird for identification.

Using AI to isolate birds in the recordings from Xeno-carto, maybe find some references or describe how a model could look/function. The model could be trained on the recordings from Xeno-Carto. Furthermore, the sound files could be checked for recording level and normalised so all sounds are in the same loudness range.

7.2 FINITO

Part I

APPENDIX

A PROJECT WORK ON MORE THAN HUMAN-CENTRED DESIGN

A.1 THE REPORT FROM THE PROJECT WORK

. **Link to a GitLab project containing the report Link:** https://gitlab.au.dk/pigeon-project/exploring-multi-species-perspectives-in-hci-through-th

A.2 THE EARLY PROTOTYPE FOR URBAN ECHOES

Link to GitLab containing the source code for the project prototype Link: https://gitlab.au.dk/pigeon-project/audioapp

B GITHUB AND CODE

B.1 MULTI PROVIDER SETUP

```
Widget build(BuildContext context) {
  return MultiProvider(
    providers: [
      // State managers
      ChangeNotifierProvider<PageStateManager>(
        create: (context) => PageStateManager(),
      ),
      ChangeNotifierProvider<MapStateManager>(
        create: (context) => MapStateManager(),
      ),
      ChangeNotifierProvider<NavigationProvider>(
        create: (context) => NavigationProvider(),
      ),
      ChangeNotifierProvider<UploadNotificationService>(
        create: (context) => UploadNotificationService(),
      ),
      Provider<DatabaseService>(
        create: (context) => DatabaseService(),
        lazy: false,
      ),
      Provider<ObservationService>(
        create: (context) => ObservationService(
          apiUrl: dotenv.env['API_URL'] ?? 'https://api.
              urbanechoes.org',
        ),
      ),
      ListenableProvider<ObservationUploader>(
        create: (context) => ObservationUploader(
          databaseService: Provider.of<DatabaseService>(context,
              listen: false),
          observationService: Provider.of<ObservationService>(
              context, listen: false),
          notificationService: Provider.of<
              UploadNotificationService>(context, listen: false),
        ),
        dispose: (context, uploader) {
          debugPrint('ObservationUploader provider dispose called
               - keeping uploader active');
        },
      ),
      ChangeNotifierProvider<RecordingPlayerService>(
```

```
37          create: (_) => RecordingPlayerService(debugMode:
                debugMode),
38          lazy: true,
39        ),
40        ChangeNotifierProvider<SpeechRecognitionService>(
41          create: (_) => SpeechRecognitionService(debugMode:
                debugMode),
42          lazy: true,
43        ),
44        ChangeNotifierProvider<BirdRecognitionService>(
45          create: (_) => BirdRecognitionService(debugMode:
                debugMode),
46          lazy: true,
47        ),
48        ChangeNotifierProvider<WordRecognitionService>(
49          create: (_) => WordRecognitionService(debugMode:
                debugMode),
50          lazy: true,
51        ),
52        ChangeNotifierProxyProvider4<
53          SpeechRecognitionService,
54          BirdRecognitionService,
55          WordRecognitionService,
56          RecordingPlayerService,
57          SpeechCoordinator>(
58          create: (context) => SpeechCoordinator(
59            speechService: Provider.of<SpeechRecognitionService>(
                  context, listen: false),
60            birdService: Provider.of<BirdRecognitionService>(
                  context, listen: false),
61            wordService: Provider.of<WordRecognitionService>(
                  context, listen: false),
62            audioService: Provider.of<RecordingPlayerService>(
                  context, listen: false),
63            observationUploader: Provider.of<ObservationUploader>(
                  context, listen: false),
64            debugMode: debugMode,
65          ),
66          update: (context, speechService, birdService, wordService
                , audioService, previous) {
67            if (previous == null) {
68              return SpeechCoordinator(
69                speechService: speechService,
70                birdService: birdService,
71                wordService: wordService,
72                audioService: audioService,
73                observationUploader: Provider.of<
                      ObservationUploader>(context, listen: false),
74                debugMode: debugMode,
75              );
76            }
77            previous.updateServices(
```

```
78              speechService: speechService,
79              birdService: birdService,
80              wordService: wordService,
81              audioService: audioService
82            );
83            return previous;
84          },
85        ),
86        ChangeNotifierProvider<SeasonService>(
87          create: (_) => SeasonService(),
88        ),
89        Provider<AppStartupService>(
90          create: (_) => AppStartupService(),
91        ),
92        Provider<bool>.value(value: debugMode),
93        ChangeNotifierProvider<LocationService>.value(value:
              locationService),
94      ],
95      child: MaterialApp(
96        title: 'Urban Echoes',
97        theme: ThemeData(
98          primarySwatch: Colors.blue,
99          bottomNavigationBarTheme: const
                BottomNavigationBarThemeData(
100           backgroundColor: Colors.black,
101           selectedItemColor: Colors.white,
102           unselectedItemColor: Colors.grey,
103         ),
104       ),
105       home: const InitialScreen(),
106     ),
107   );
108 }
```

Listing B.1: MultiProvider setup in Flutter

B.2 FASTAPI BACKEND

.1 FASTAPI BACKEND CODE

```
1 from fastapi import FastAPI, HTTPException, Query, Depends
2 import requests
3 import psycopg2
4 import random
5 import os
6 import logging
7
8 from fastapi.middleware.cors import CORSMiddleware
9 from dotenv import load_dotenv
```

```
from psycopg2.extras import RealDictCursor

load_dotenv()

app = FastAPI()

app.add_middleware(
    CORSMiddleware,
    allow_origins=["*"],  # Change to specific domains for security
    allow_credentials=True,
    allow_methods=["*"],
    allow_headers=["*"],
)

# Configure logging
logging.basicConfig(level=logging.INFO)
logger = logging.getLogger(__name__)

# Database connection function
def get_db_connection():
    return psycopg2.connect(
        user=os.getenv("DB_USER"),
        password=os.getenv("DB_PASSWORD"),
        host=os.getenv("DB_HOST"),
        port=os.getenv("DB_PORT", 5432),
        database=os.getenv("DB_NAME ", "urban_echoes_db "),
        sslmode="require"
    )

EBIRD_API_URL = "https://api.ebird.org/v2/data/obs/geo/recent"
EBIRD_TAXONOMY_URL = "https://api.ebird.org/v2/ref/taxonomy/ebird"
XENO_CANTO_API = "https://www.xeno-canto.org/api/2/recordings"
EBIRD_API_KEY = os.getenv("EBIRD_API_KEY")
DATABASE_URL = os.getenv("DATABASE_URL")

if not EBIRD_API_KEY:
    raise ValueError("EBIRD_API_KEY is missing! Set it in Azure.")

if not DATABASE_URL:
    raise ValueError("DATABASE_URL is missing! Set it in Azure.")

LAT = 56.2639 # Copenhagen coordinates TODO change to your location
LON = 9.5018  # Copenhagen coordinates TODO change to your location

async def get_danish_taxonomy():
    """Fetch the eBird taxonomy with Danish names."""
    headers = {"X-eBirdApiToken": EBIRD_API_KEY}
```

```
    params = {"fmt": "json", "locale": "da"}

    try:
        response = requests.get(EBIRD_TAXONOMY_URL, headers=
            headers, params=params)
        response.raise_for_status()
        taxonomy_data = response.json()
        return {species["speciesCode"]: species["comName"] for
            species in taxonomy_data}
    except requests.exceptions.RequestException as e:
        raise HTTPException(status_code=500, detail=f"Error
            fetching taxonomy: {str(e)}")

@app.get("/observations")
def get_observations(after_timestamp: str = Query(None,
    description="Fetch only observations created after this
    timestamp")):
    """Fetch bird observations from the database, with optional
        filtering by timestamp."""
    try:
        conn = get_db_connection()
        cursor = conn.cursor(cursor_factory=RealDictCursor)

        if after_timestamp:
            cursor.execute("""
                SELECT id, bird_name, scientific_name,
                    sound_directory, latitude, longitude,
                       observation_date, observation_time,
                           observer_id, created_at,
                       quantity, is_test_data, test_batch_id
                FROM bird_observations
                WHERE created_at > %s
                ORDER BY created_at ASC
            """, (after_timestamp,))
            logger.info(f"Fetching observations created after {
                after_timestamp}")
        else:
            cursor.execute("""
                SELECT id, bird_name, scientific_name,
                    sound_directory, latitude, longitude,
                       observation_date, observation_time,
                           observer_id, created_at,
                       quantity, is_test_data, test_batch_id
                FROM bird_observations
                ORDER BY created_at ASC
            """)
            logger.info("Fetching all observations")

        observations = cursor.fetchall()
        logger.info(f"Returning {len(observations)} observations"
            )

```

```
        cursor.close()
        conn.close()

        return {"observations": observations}
    except Exception as e:
        logger.error(f"Error fetching observations: {str(e)}")
        raise HTTPException(status_code=500, detail=f"Internal
            Server Error: {str(e)}")

@app.get("/birds")
def get_birds():
    try:
        conn = get_db_connection()
        cursor = conn.cursor(cursor_factory=RealDictCursor)
        cursor.execute("SELECT common_name, scientific_name,
            danish_name FROM birds")
        birds = cursor.fetchall()
        cursor.close()
        conn.close()
        return {"birds": birds}
    except Exception as e:
        logger.error(f"Error fetching birds: {str(e)}")
        raise HTTPException(status_code=500, detail="Internal
            Server Error")

@app.get("/birdsound")
def get_bird_sound(scientific_name: str):
    params = {"query": scientific_name}
    response = requests.get(XENO_CANTO_API, params=params)

    if response.status_code != 200:
        return {"error": "Failed to fetch recordings"}

    data = response.json()
    recordings = data.get("recordings", [])

    if not recordings:
        return {"error": "No recordings found"}

    high_quality = [rec for rec in recordings if rec.get("q") in
        ["A", "B"]]

    if not high_quality:
        return {"error": "No high-quality recordings available"}

    selected = random.choice(high_quality)
    sound_url = f"https://www.xeno-canto.org/{selected['id']}/
        download"

    return sound_url

@app.get("/search_birds")
```

```python
def search_birds(query: str = Query(..., min_length=1, description="Bird search query")):
    """Search birds by Danish name dynamically"""
    try:
        conn = get_db_connection()
        cursor = conn.cursor()
        cursor.execute("""
            SELECT common_name, scientific_name
            FROM birds
            WHERE common_name ILIKE %s
            LIMIT 10
        """, (f"%{query}%",))
        birds = [{"common_name": row[0], "scientificName": row[1]} for row in cursor.fetchall()]
        cursor.close()
        conn.close()
        return {"birds": birds}
    except Exception as e:
        logger.error(f"Error searching birds: {str(e)}")
        raise HTTPException(status_code=500, detail="Internal Server Error")

@app.get("/birdsOLD")
async def get_bird_list():
    """Fetch recent bird observations with Danish names and corresponding sounds."""
    headers = {"X-eBirdApiToken": EBIRD_API_KEY}
    params = {
        "lat": LAT,
        "lng": LON,
        "fmt": "json",
        "maxResults": 100,
        "includeProvisional": True
    }

    try:
        danish_names = await get_danish_taxonomy()
        response = requests.get(EBIRD_API_URL, headers=headers, params=params)
        response.raise_for_status()
        bird_data = response.json()

        birds = []
        for bird in bird_data:
            species_code = bird.get("speciesCode")
            scientific_name = bird.get("sciName")

            bird_info = {
                "danishName": danish_names.get(species_code, bird.get("comName")),
                "scientificName": scientific_name,
                "observationDate": bird.get("obsDt"),
```

```
                "location": bird.get("locName"),
                "speciesCode": species_code
            }
            birds.append(bird_info)

        return {
            "birds": birds,
            "count": len(birds),
            "location": f"Coordinates: {LAT}, {LON}"
        }
    except requests.exceptions.RequestException as e:
        raise HTTPException(status_code=500, detail=f"Error
            fetching bird data: {str(e)}")

@app.get("/health")
async def health_check():
    """Health check endpoint to verify that the API is running.
        """
    return {"status": "ok"}
```

Listing 2: Full FastAPI Backend Code

A EVALUATIONS

A.1 DIARY STUDY RESPONSES

Link to Google Docs folder containing the evaluation diary responses Link: https://drive.google.com/drive/folders/1lUAUgP384W6d9hCbFUsGoeR75pICr5
usp=sharing

BIBLIOGRAPHY

[1] William C Adams. "Conducting semi-structured interviews." In: *Handbook of practical program evaluation* (2015), pp. 492–505.

[2] American Birding Association. *ABA Code of Birding Ethics*. https://www.aba.org/aba-code-of-birding-ethics/. Accessed: 2025-04-10.

[3] Rochester Birding Association. *Birdsong*. https://rochesterbirding.org/conservation/ethical-birding/birdsong/. Accessed: 2025-04-10.

[4] Baseflow. *Geolocator: A Flutter plugin for accessing geolocation information*. https://pub.dev/packages/geolocator. Accessed: 2025-05-16. 2024.

[5] Heidi R Biggs, Jeffrey Bardzell, and Shaowen Bardzell. "Watching myself watching birds: Abjection, ecological thinking, and posthuman design." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–16.

[6] Eli Blevis. "Sustainable interaction design: invention & disposal, renewal & reuse." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2007, pp. 503–512.

[7] Sophie Boehm, Louise Jeffery, Judit Hecke, Clea Schumer, Joel Jaeger, Claire Fyson, Kelly Levin, Anna Nilsson, Stephen Naimoli, Emily Daly, et al. "State of climate action 2023." In: (2022).

[8] Christina Bremer, Bran Knowles, and Adrian Friday. "Have We Taken On Too Much?: A Critical Review of the Sustainable HCI Landscape." In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. ISBN: 9781450391573. DOI: 10.1145/3491102.3517609. URL: https://doi.org/10.1145/3491102.3517609.

[9] Janet Cardiff. *Forest Walk*. Audio walk created during a residency at the Banff Centre for the Arts. 1991. URL: https://cardiffmiller.com/walks/forest-walk/.

[10] Janet Cardiff. *Her Long Black Hair*. https://cardiffmiller.com/walks/her-long-black-hair/. Audio walk with photographs, 46 minutes. Curated by Tom Eccles for the Public Art Fund, June 17–September 13, 2004, Central Park, New York. 2004.

[11] Louise Chawla. "Significant life experiences revisited: A review of research on sources of environmental sensitivity." In: *The Journal of environmental education* 29.3 (1998), pp. 11–21.

[12] Jeffrey P Cohn. "Citizen science: Can volunteers do real research?" In: *BioScience* 58.3 (2008), pp. 192–197.

[13] FleaFlet Contributors. *FleaFlet: A Flutter mapping library inspired by Leaflet.js*. https://docs.fleaflet.dev/. Accessed: 2025-05-16. 2024.

[14] CSDC Corp. *speech_to_text: A Flutter plugin that exposes device specific speech to text recognition capability*. https://pub.dev/packages/speech_to_text. Version 7.0.0, Accessed: 2025-05-14. 2024.

[15] Paul Coulton and Joseph Galen Lindley. "More-than human centred design: Considering other things." In: *The Design Journal* 22.4 (2019), pp. 463–481.

[16] Stuart Cunningham, Iain McGregor, Jonathan Weinel, John Darby, and Tony Stockman. "Towards a Framework of Aesthetics in Sonic Interaction." In: *Proceedings of the 18th International Audio Mostly Conference*. AM '23. Edinburgh, United Kingdom: Association for Computing Machinery, 2023, 109–115. ISBN: 9798400708183. DOI: 10.1145/3616195.3616219. URL: https://doi.org/10.1145/3616195.3616219.

[17] Dansk Ornitologisk Forening. *DOFbasen - Dansk Ornitologisk Forening*. Accessed: 2025-03-31. 2025. URL: https://dofbasen.dk/.

[18] Cedric Donnat and Flutter Community. *flutter_tts: Flutter Text to Speech Plugin*. https://pub.dev/packages/flutter_tts. Accessed: 2025-05-01. 2024.

[19] Robert R. Dunn, Michael C. Gavin, Monica C. Sanchez, and Jennifer N. Solomon. "The Pigeon Paradox: Dependence of Global Conservation on Urban Nature." In: *Conservation Biology* 20.6 (2006), pp. 1814–1816. ISSN: 08888892, 15231739. URL: http://www.jstor.org/stable/4124710 (visited on 11/08/2024).

[20] ElevenLabs Inc. *ElevenLabs: Free Text to Speech & AI Voice Generator*. Accessed: 2025-04-22. 2025. URL: https://elevenlabs.io/.

[21] Jixiang Fan, Derek Haqq, Morva Saaty, Wei-Lu Wang, and Scott McCrickard. "Diary Study as an Educational Tool: An Experience Report from an HCI Course." In: *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1*. 2025, pp. 311–317.

[22] Flutter Team. *Flutter - Build apps for any screen*. Accessed: 2025-04-21. 2025. URL: https://flutter.dev/.

[23] Xeno canto Foundation and Contributors. *Xeno-canto: Sharing bird sounds from around the world*. Accessed: 2024-03-06. 2024. URL: https://xeno-canto.org.

[24] William Gaver, Andy Boucher, Michail Vanis, Andy Sheen, Dean Brown, Liliana Ovalle, Naho Matsuda, Amina Abbas-Nazari, and Robert Phillips. "My naturewatch camera: disseminating practice research with a cheap and easy DIY design." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, pp. 1–13.

[25] Lon Åke Erni Johannes Hansson, Teresa Cerratto Pargman, and Daniel Sapiens Pargman. "A Decade of Sustainable HCI: Connecting SHCI to the Sustainable Development Goals." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: 10.1145/3411764.3445069. URL: https://doi.org/10.1145/3411764.3445069.

[26] J Berton C Harris and David G Haskell. "Simulated birdwatchers' playback affects the behavior of two tropical birds." In: *PLoS One* 8.10 (2013), e77902.

[27] Alan Irwin. *Citizen science: A study of people, expertise and sustainable development*. Routledge, 2002.

[28] Mari Laakso and L Tiina Sarjakoski. "Sonic maps for hiking—use of sound in enhancing the map use experience." In: *The Cartographic Journal* 47.4 (2010), pp. 300–307.

[29] David Lack. "The breeding seasons of European birds." In: *Ibis* 92.2 (1950), pp. 288–316.

[30] Mark Lawton, Stuart Cunningham, and Ian Convery. "Nature soundscapes: an audio augmented reality experience." In: *Proceedings of the 15th International Audio Mostly Conference*. 2020, pp. 85–92.

[31] Jiang Liu, Jian Kang, Holger Behm, and Tao Luo. "Effects of landscape on soundscape perception: Soundwalks in city parks." In: *Landscape and urban planning* 123 (2014), pp. 30–40.

[32] Szu-Yu Liu, Jen Liu, Kristin Dew, Patrycja Zdziarska, Maya Livio, and Shaowen Bardzell. "Exploring noticing as method in design research." In: *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*. 2019, pp. 377–380.

[33] Jennifer C Mankoff, Eli Blevis, Alan Borning, Batya Friedman, Susan R Fussell, Jay Hasbrouck, Allison Woodruff, and Phoebe Sengers. "Environmental sustainability and interaction." In: *CHI'07 extended abstracts on Human factors in computing systems*. 2007, pp. 2121–2124.

[34] Mark Monmonier. "Rhumb lines and map wars: A social history of the Mercator projection." In: *Rhumb Lines and Map Wars*. University of Chicago Press, 2010.

[35] Stuart Moran, Nadia Pantidi, Tom Rodden, Alan Chamberlain, Chloe Griffiths, Davide Zilli, Geoff Merrett, and Alex Rogers. "Listening to the forest and its curators: lessons learnt from a bioacoustic smartphone application deployment." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, 2387–2396. ISBN: 9781450324731. DOI: 10.1145/2556288.2557022. URL: https://doi.org/10.1145/2556288.2557022.

[36] Netfugl. *DK listen*. Accessed: 2025-05-14. 2025. URL: https://www.netfugl.dk/artslister/dk.

[37] Luan Nico and Flutter Community. *audioplayers*. https://pub.dev/packages/audioplayers. Accessed: 2025-05-16. 2024.

[38] Valentina Nisi, Enrico Costanza, and Mara Dionisio. "Placing Location-Based Narratives in Context Through a Narrator and Visual Markers." In: *Interacting with Computers* 29.3 (July 2016), pp. 287–305. ISSN: 0953-5438. DOI: 10.1093/iwc/iww020. eprint: https://academic.oup.com/iwc/article-pdf/29/3/287/11149005/iww020.pdf. URL: https://doi.org/10.1093/iwc/iww020.

[39] Claudia Núñez-Pacheco and Anton Poikolainen Rosén. "Articulating felt senses for more-than-human design: A viewpoint for noticing." In: *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2024, pp. 1029–1043.

[40] Cornell Lab of Ornithology. *Merlin Bird ID App*. 2025. URL: https://merlin.allaboutbirds.org/.

[41] Tilde Pedersen, Edith Terte Andersen, and Anders Sundnes Løvlie. "Designing a "no interface" audio walk." In: *Museums and the Web*. 2019.

[42] Marianne Graves Petersen, Ole Sejer Iversen, Peter Gall Krogh, and Martin Ludvigsen. "Aesthetic interaction: a pragmatist's aesthetics of interactive systems." In: *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. 2004, pp. 269–276.

[43] Manon Raimbault and Danièle Dubois. "Urban soundscapes: Experiences and knowledge." In: *Cities* 22.5 (2005), pp. 339–350. ISSN: 0264-2751. DOI: https://doi.org/10.1016/j.cities.2005.05.003. URL: https://www.sciencedirect.com/science/article/pii/S0264275105000557.

[44] Eleanor Ratcliffe, Birgitta Gatersleben, and Paul T. Sowden. "Bird sounds and their contributions to perceived attention restoration and stress recovery." In: *Journal of Environmental Psychology* 36 (2013), pp. 221–228. ISSN: 0272-4944. DOI: https://doi.org/10.

1016/j.jenvp.2013.08.004. URL: https://www.sciencedirect.com/science/article/pii/S0272494413000650.

[45] Chandler S Robbins. "Effect of time of day on bird activity." In: *Studies in avian biology* 6.3 (1981), pp. 275–286.

[46] Anton Poikolainen Rosén, Maria Normark, and Mikael Wiberg. "Noticing the environment–A design ethnography of urban farming." In: *Nordic Human-Computer Interaction Conference*. 2022, pp. 1–13.

[47] R Murray Schafer. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster, 1993.

[48] Simone Scully. *Play a Song, Hurt a Bird?* https://www.audubon.org/news/play-song-hurt-bird. Accessed: 2025-04-10. 2013.

[49] David Sibley. *The Proper Use of Playback in Birding*. https://www.sibleyguides.com/2011/04/the-proper-use-of-playback-in-birding/. Accessed: 2025-04-10. 2011.

[50] Tore Slagsvold. "Bird song activity in relation to breeding cycle, spring weather, and environmental phenology." In: *Ornis Scandinavica* (1977), pp. 197–222.

[51] Brian L. Sullivan, Christopher L. Wood, Marshall J. Iliff, Richard E. Bonney, Daniel Fink, and Steve Kelling. "eBird: A citizen-based bird observation network in the biological sciences." In: *Biological Conservation* 142 (2009), pp. 2282–2292.

[52] Alex Taylor. "After interaction." In: *Interactions* 22.5 (Aug. 2015), 48–53. ISSN: 1072-5520. DOI: 10.1145/2809888. URL: https://doi.org/10.1145/2809888.

[53] Samuel Thulin. "Sound maps matter: expanding cartophony." In: *Social & Cultural Geography* 19.2 (2018), pp. 192–210. DOI: 10.1080/14649365.2016.1266028. eprint: https://doi.org/10.1080/14649365.2016.1266028. URL: https://doi.org/10.1080/14649365.2016.1266028.

[54] Anna Lowenhaupt Tsing. *The mushroom at the end of the world: On the possibility of life in capitalist ruins*. Princeton University Press, 2015.

[55] United Nations, Department of Economic and Social Affairs, Population Division. *2018 Revision of World Urbanization Prospects*. Accessed: 2025-03-17. 2018. URL: https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html.

[56] Yang Wang, Jiaogen Zhou, Caiyun Zhang, Zhaopeng Luo, Xuexue Han, Yanzhu Ji, and Jihong Guan. "Bird object detection: Dataset construction, model performance evaluation, and model lightweighting." In: *Animals* 13.18 (2023), p. 2924.

[57] Kendall Wrightson. "An introduction to acoustic ecology." In: *Soundscape: The journal of acoustic ecology* 1.1 (2000), pp. 10–13.

[58] eBird. *eBird Live Subscriptions*. Accessed: 2025-03-31. 2025. URL: https://ebird.org/livesubs.

[59] iNaturalist. *iNaturalist: A Community for Naturalists*. Accessed: April 8, 2025. 2025. URL: https://www.inaturalist.org/.