Leaflet: Transcription Software Reimagined for Journalists

[Redacted], [Redacted], and Grace Zhou

With the advent of mobile technology, most modern interviews exist not solely in handwritten notes and human memory, but rather in audio files. To aid in the writing process, many reporters use automated transcription software in addition to handwritten notes taken during the interview to capture exact or near-exact wording and ambiance. While transcription software exists, it is fragmented from the journalist work flow, as it solves only one of many moving parts during the interview and writing process.

With the existing challenges in mind, we made **Leaflet**, a web platform that transcribes and beyond. Journalists can attach news articles, listen to precise audio segments, and bridge digital notes and transcriptions with their workspace for storycraft. In creating and unifying features that translate the physical into the digital, as well as interactions across digital components, we optimize the journalist's workflow.

Qualitative Research

Before creating the application, we studied five popular transcription tools. We input real audio from interviews to gauge their accuracy and available features. Nearly every tool provided a correction interface for the transcript, identified different speakers, and showed timestamps. However, none had features that specifically targeted the reporting process – the tools' capabilities were general in order to please wider audiences. As a result, they didn't help users with much beyond transcribing audio, which is often just one step of a user's workflow.

	Transcribe. wreally	**Otter.ai**	**Transcribe for iPhone**	**Trint**	**Rev**
Records within app	no	yes	no	yes	yes
Correct transcript	yes	yes	yes	yes	yes
Collaboration/sharing	no	yes	yes	yes	yes
Records/transcribes phone calls	no	yes	no	yes	yes
Click words to play back sections	no	yes	no	yes	yes
Speaker matching	yes	yes	no	yes	yes
Topic identification	no	yes	no	no	no
Search across all transcripts	no	yes	yes	no	yes
Timestamps	yes	yes	yes	yes	yes
Highlight quotes	no	yes	no	yes	yes
Transcription accuracy	9/10	9/10	8.5/10	8.7/10	9.9/10 human, 8/10 machine

A comparison of various transcription tools' features

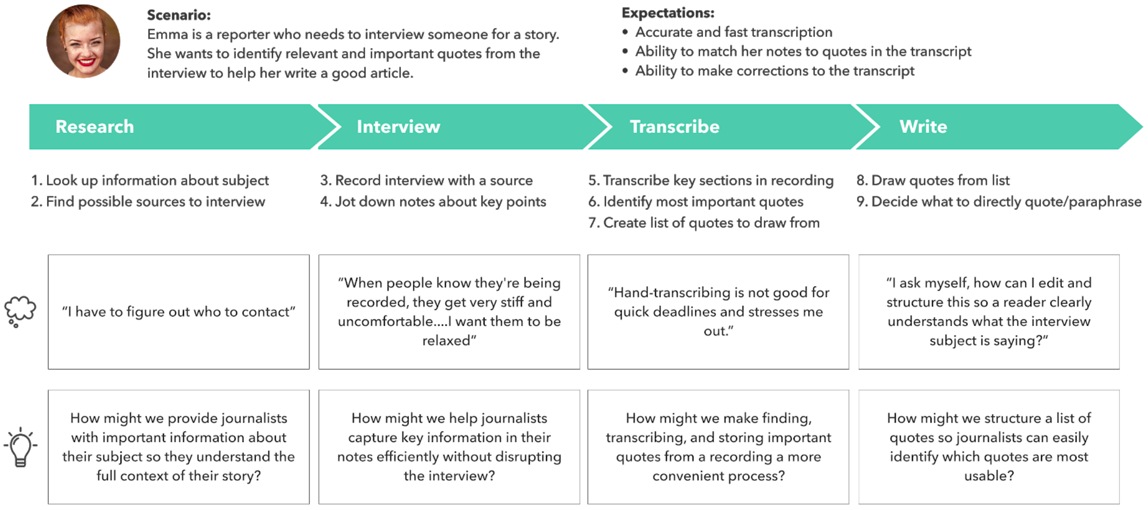
Throughout the quarter, we spoke to local reporters at *The Stanford Daily* and industry reporters. We gained greater insight into the nuances of their workflow by asking how they record audio and how often, what existing affordances and drawbacks they saw in transcription tools, when and how they transfer information to a text editor (often Google Docs), and other questions.

We learned that before doing an interview or attending an event, reporters conduct research to gain context about their subject. While recording audio at these events, they limit the presence of electronic devices to avoid distracting others. They also take handwritten notes to make note of important points, often including timestamps. Later, they use these notes to find key phrases and quotes within the transcript to use for their story. They store these quote in their own makeshift "quote banks," and some desire a better method of saving sections of the transcript for later. When asked their opinions on being able to link information such as Wikipedia pages to transcripts, they said it was not necessary to them and that they cared most about the quality of the transcription. However, the exception was that they wanted to be able to find and link similar news articles to their transcripts.

The reporters we spoke to cared more about improving the way they navigate transcripts and store key insights digitally than improving how they record audio. This led to our shift away from technical solutions that would be intrusive to the interviews themselves, but rather complement and streamline the identification of pertinent information.

One industry reporter's responses stood out to us: unlike the student reporters we interviewed, she never uses transcription tools because transcribing by hand helps familiarize her with the material. Concerned about privacy and lawsuits, she also rarely records interviews and regularly throws her notes away. Because security and policy were out of the scope of this project, we realized we should target our tool to student and early-career journalists who are open to trying new technologies and may interview fewer anonymous people.

With the insights from our interviews, we mapped out a reporter's typical writing process to identify opportunities for our tool to help.



Scenario:
Emma is a reporter who needs to interview someone for a story. She wants to identify relevant and important quotes from the interview to help her write a good article.

Expectations:
- Accurate and fast transcription
- Ability to match her notes to quotes in the transcript
- Ability to make corrections to the transcript

	Research	Interview	Transcribe	Write
	1. Look up information about subject 2. Find possible sources to interview	3. Record interview with a source 4. Jot down notes about key points	5. Transcribe key sections in recording 6. Identify most important quotes 7. Create list of quotes to draw from	8. Draw quotes from list 9. Decide what to directly quote/paraphrase
	"I have to figure out who to contact"	"When people know they're being recorded, they get very stiff and uncomfortable....I want them to be relaxed"	"Hand-transcribing is not good for quick deadlines and stresses me out."	"I ask myself, how can I edit and structure this so a reader clearly understands what the interview subject is saying?"
	How might we provide journalists with important information about their subject so they understand the full context of their story?	How might we help journalists capture key information in their notes efficiently without disrupting the interview?	How might we make finding, transcribing, and storing important quotes from a recording a more convenient process?	How might we structure a list of quotes so journalists can easily identify which quotes are most usable?

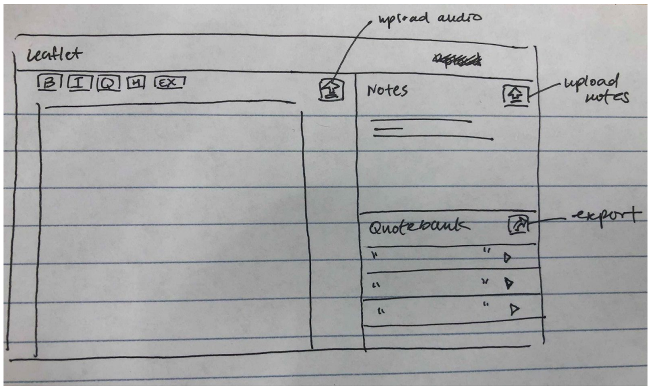
Ideating and Prototyping

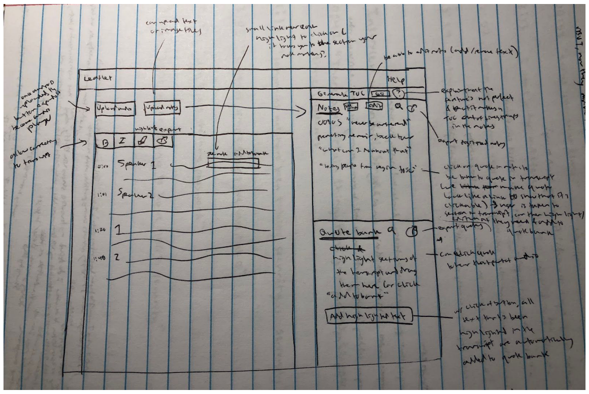
After learning about journalists' transcription needs and processes, we brainstormed and discussed ideas for features we thought would be useful for a reporter before, during, and after transcribing. Beyond the main functionality of transcribing audio, we decided to focus on these features:

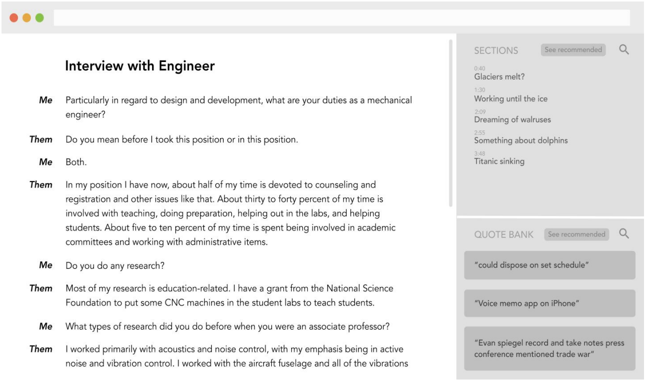
- Quote bank for saving and storing notable quotes
- News search for finding related published stories
- Correctional interface for editing the automated transcript
- Computer vision for converting images of handwritten notes to digitized text

We considered many other ideas but ultimately moved away from them because of time and technical constraints. For instance, we were so excited with the idea of converting handwritten notes into digitally formatted text that we brainstormed shorthand symbols that could automatically highlight, paragraph, and indent text. We thought of creating a Google Docs plugin that did this. However, we realized that time in this direction would detract from the implementation of fundamental transcription interactivity such as audio-play upon text selection, which we found to be more of an immediate need. Thus, we decided to implement the basics of note conversion and focus on the core functionality of audio-to-text.

We created sketches and prototypes to get an idea of the final application's visual structure and technical requirements.






Interview with Engineer

Me Particularly in regard to design and development, what are your duties as a mechanical engineer?

Them Do you mean before I took this position or in this position.

Me Both.

Them In my position I have now, about half of my time is devoted to counseling and registration and other issues like that. About thirty to forty percent of my time is involved with teaching, doing preparation, helping out in the labs, and helping students. About five to ten percent of my time is spent being involved in academic committees and working with administrative items.

Me Do you do any research?

Them Most of my research is education-related. I have a grant from the National Science Foundation to put some CNC machines in the student labs to teach students.

Me What types of research did you do before when you were an associate professor?

Them I worked primarily with acoustics and noise control, with my emphasis being in active noise and vibration control. I worked with the aircraft fuselage and all of the vibrations

SECTIONS See recommended

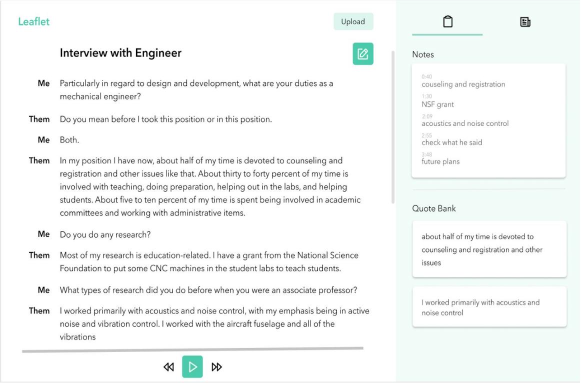
0:40 Glaciers melt?
1:30 Working until the ice
2:09 Dreaming of walruses
2:55 Something about dolphins
3:48 Titanic sinking

QUOTE BANK See recommended

"could dispose on set schedule"

"Voice memo app on iPhone"

"Evan spiegel record and take notes press conference mentioned trade war"


Leaflet
Upload

Interview with Engineer

Me Particularly in regard to design and development, what are your duties as a mechanical engineer?

Them Do you mean before I took this position or in this position.

Me Both.

Them In my position I have now, about half of my time is devoted to counseling and registration and other issues like that. About thirty to forty percent of my time is involved with teaching, doing preparation, helping out in the labs, and helping students. About five to ten percent of my time is spent being involved in academic committees and working with administrative items.

Me Do you do any research?

Them Most of my research is education-related. I have a grant from the National Science Foundation to put some CNC machines in the student labs to teach students.

Me What types of research did you do before when you were an associate professor?

Them I worked primarily with acoustics and noise control, with my emphasis being in active noise and vibration control. I worked with the aircraft fuselage and all of the vibrations

Notes

0:40 couseling and registration
1:30 NSF grant
2:09 acoustics and noise control
2:55 check what he said
3:48 future plans

Quote Bank

about half of my time is devoted to counseling and registration and other issues

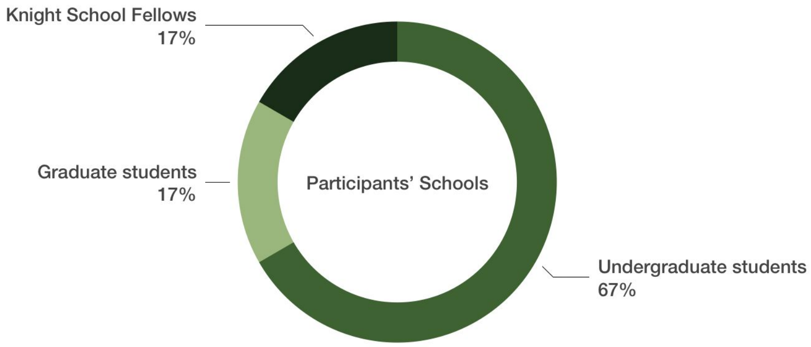
I worked primarily with acoustics and noise control

Feedback from Journalists

While prototyping and developing, it was critical for us to continue talking to current journalists about their transcription processes and their thoughts on different components of Leaflet. We did this in two main ways: surveying and testing.

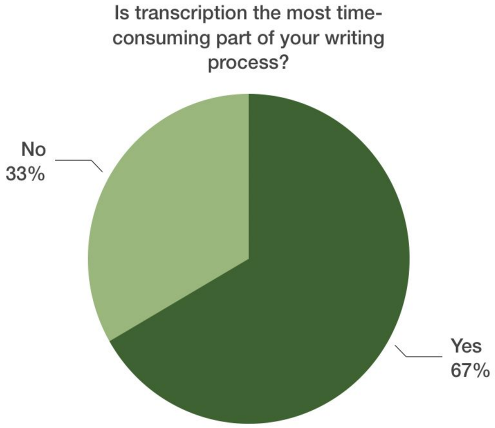
Survey

We sent a Google Form to the entire staff of *The Stanford Daily* and to the current class of John S. Knight Fellows. We received responses primarily from undergraduate student journalists.

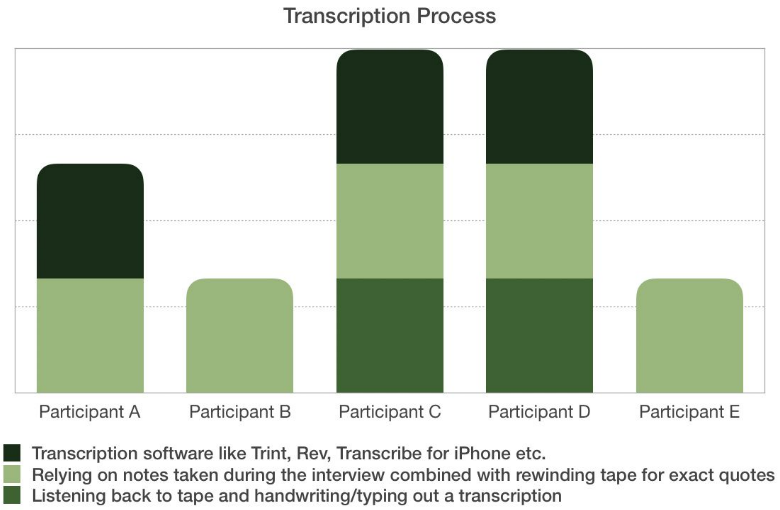


Within that group, those who responded were primarily sophomores, an important subset as they likely have at least a year of student journalism under their belt and have developed a strategy and workflow for interviewing, transcribing, and writing.

We also asked survey responders to categorize their describe transcription process. From this we learned that a majority of them currently use automated transcription software and, thus, wouldn't find the shift to Leaflet out of the ordinary. We also learned that all of them rewind their tapes and listen back to collect exact quotes for their stories, a process that is extremely time-consuming – as evidenced below – and inefficient. The quote bank feature of Leaflet would completely eliminate the need to do this.



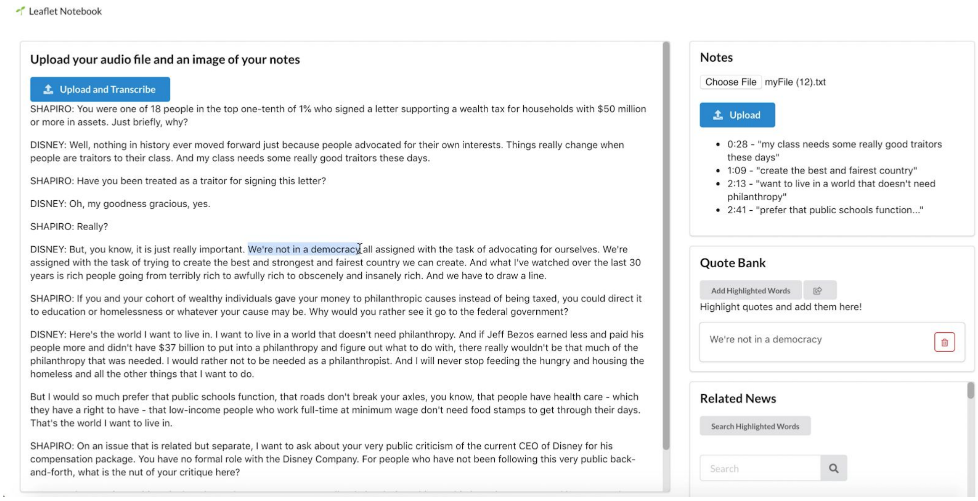
When presented with a preliminary prototype of the current quote bank feature, 100% of participants said that it would be useful and they would use it in addition to their current tools, with 67% saying they would use it instead of their current tools.



All participants also take notes during interviews. With the capacity to translate handwritten, in-interview notes into text tagged with clickable timestamps, Leaflet would be built to allow these journalists to continue that practice, making it simpler and easier to integrate it with the rest of the writing process. Eighty-three percent of participants said that a digital note-taking feature would be useful and that they would use it in addition to their current tools.

Tests

During development, we tested Leaflet on members of our target demographic: young journalists. We asked them to complete a series of tasks while expressing their thoughts out loud. We also asked them questions about their experiences with transcribing audio. This helped us understand what aspects of the tool were intuitive and useful and what could be improved.



SHAPIRO: You were one of 18 people in the top one-tenth of 1% who signed a letter supporting a wealth tax for households with $50 million or more in assets. Just briefly, why?

DISNEY: Well, nothing in history ever moved forward just because people advocated for their own interests. Things really change when people are traitors to their class. And my class needs some really good traitors these days.

SHAPIRO: Have you been treated as a traitor for signing this letter?

DISNEY: Oh, my goodness gracious, yes.

SHAPIRO: Really?

DISNEY: But, you know, it is just really important. We're not in a democracy all assigned with the task of advocating for ourselves. We're assigned with the task of trying to create the best and strongest and fairest country we can create. And what I've watched over the last 30 years is rich people going from terribly rich to awfully rich to obscenely and insanely rich. And we have to draw a line.

SHAPIRO: If you and your cohort of wealthy individuals gave your money to philanthropic causes instead of being taxed, you could direct it to education or homelessness or whatever your cause may be. Why would you rather see it go to the federal government?

DISNEY: Here's the world I want to live in. I want to live in a world that doesn't need philanthropy. And if Jeff Bezos earned less and paid his people more and didn't have $37 billion to put into a philanthropy and figure out what to do with, there really wouldn't be that much of the philanthropy that was needed. I would rather not to be needed as a philanthropist. And I will never stop feeding the hungry and housing the homeless and all the other things that I want to do.

But I would so much prefer that public schools function, that roads don't break your axles, you know, that people have health care - which they have a right to have - that low-income people who work full-time at minimum wage don't need food stamps to get through their days. That's the world I want to live in.

SHAPIRO: On an issue that is related but separate, I want to ask about your very public criticism of the current CEO of Disney for his compensation package. You have no formal role with the Disney Company. For people who have not been following this very public back-and-forth, what is the nut of your critique here?

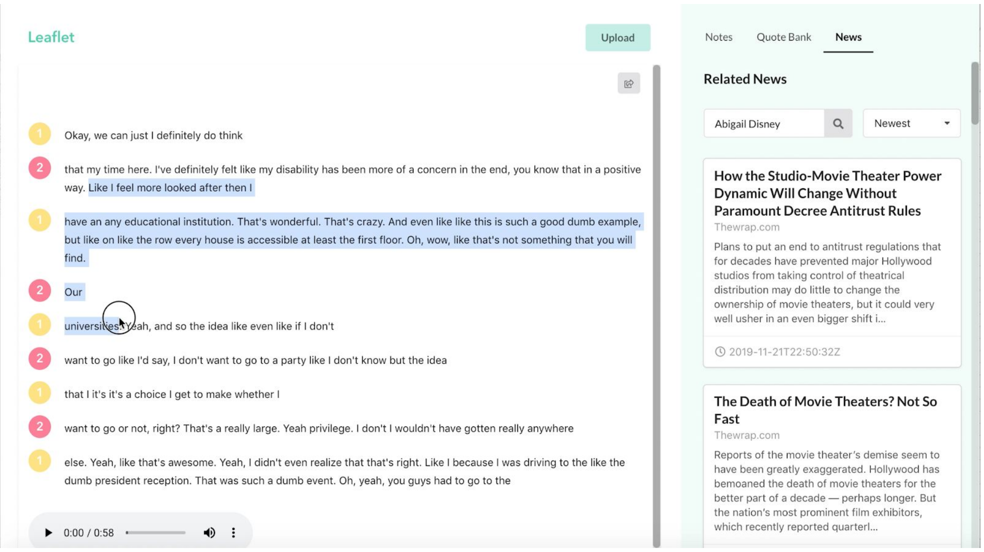
myFile (12).txt

- 0:28 - "my class needs some really good traitors these days"
- 1:09 - "create the best and fairest country"
- 2:13 - "want to live in a world that doesn't need philanthropy"
- 2:41 - "prefer that public schools function..."

Highlight quotes and add them here!

We're not in a democracy

Interface presented to users in the first round of testing



1 Okay, we can just I definitely do think

2 that my time here. I've definitely felt like my disability has been more of a concern in the end, you know that in a positive way. Like I feel more looked after then I

1 have an any educational institution. That's wonderful. That's crazy. And even like like this is such a good dumb example, but like on like the row every house is accessible at least the first floor. Oh, wow, like that's not something that you will find.

2 Our

1 universities. Yeah, and so the idea like even like if I don't

2 want to go like I'd say, I don't want to go to a party like I don't know but the idea

1 that I it's it's a choice I get to make whether I

2 want to go or not, right? That's a really large. Yeah privilege. I don't I wouldn't have gotten really anywhere

1 else. Yeah, like that's awesome. Yeah, I didn't even realize that that's right. Like I because I was driving to the like the dumb president reception. That was such a dumb event. Oh, yeah, you guys had to go to the

Related News

How the Studio-Movie Theater Power Dynamic Will Change Without Paramount Decree Antitrust Rules

Thewrap.com

Plans to put an end to antitrust regulations that for decades have prevented major Hollywood studios from taking control of theatrical distribution may do little to change the ownership of movie theaters, but it could very well usher in an even bigger shift i...

2019-11-21T22:50:32Z

The Death of Movie Theaters? Not So Fast

Thewrap.com

Reports of the movie theater's demise seem to have been greatly exaggerated. Hollywood has bemoaned the death of movie theaters for the better part of a decade — perhaps longer. But the nation's most prominent film exhibitors, which recently reported quarterl...

Interface presented in the second round

When trying the handwritten-to-digital notes conversion feature, one user became excited and commented, “A lot of times I’m forced to write game notes on a computer, but I think it’d be really helpful to take notes on paper...You can think a lot clearer that way.” Another was not sure what the “Notes” section was for at first, but thought it could be useful if they “had used this tool once already.” This suggested the feature would help journalists, but it may require more explanation because of its novelty.

Users liked the quote bank and found it easy to use. They also found the related news section “super helpful” for research. One user remarked, “Every interview that I’ve done, you have to at least explain the person. And to help put a person in context, you have to find the most recent thing they’ve done...I really like that feature a lot.”

As a result of testing, we made several changes to the tool. For example, many users expected the notes to link to the transcript in some way, so we made sure to implement this later on. We noticed that when searching for news, users gravitated towards the search box rather than the “Search Highlighted Words” button, so we took out the button to reduce clutter. Users said they would like a way to save or pin the news articles for later, so we added this functionality. We also fixed a few bugs that appeared during testing.

Developing Leaflet

Leaflet was made with a React frontend and Python backend. We also drew upon Flask, Semantic UI, Microsoft Computer Vision, Google Speech-to-Text, and News API. Although we considered several mobile features, we chose to develop a web-based application because it is easier for journalists to navigate and fits with our technical skills. We used React for the foundation of the interface and Semantic UI for visual components such as buttons, icons, and input. React is the standard for creating responsive web applications that can update without reloading the page, which is perfect for a transcription tool that requires a lot of manipulation from the user. Semantic UI is a reputable development framework that allows for some customizations, and we felt its modern, clean aesthetic was a fit for Leaflet's audience. Due to time constraints, we were unable to change the default font, font sizes, and a few other aesthetic elements to be exactly like our planned design, but Semantic UI did heavily streamline our frontend development.

That left a daunting task ahead of us: programming and manipulating the backend to curate the content we wanted to display. Computer vision for handwritten notes was the earliest feature we implemented; we weren't sure if it was possible, so we tried it immediately to resolve our uncertainty. After testing and researching different computer vision APIs that could recognize written characters, we chose Microsoft's for its low cost and relatively high accuracy. We created a Python script to communicate with the API and used Flask to pass in images users uploaded to the interface.

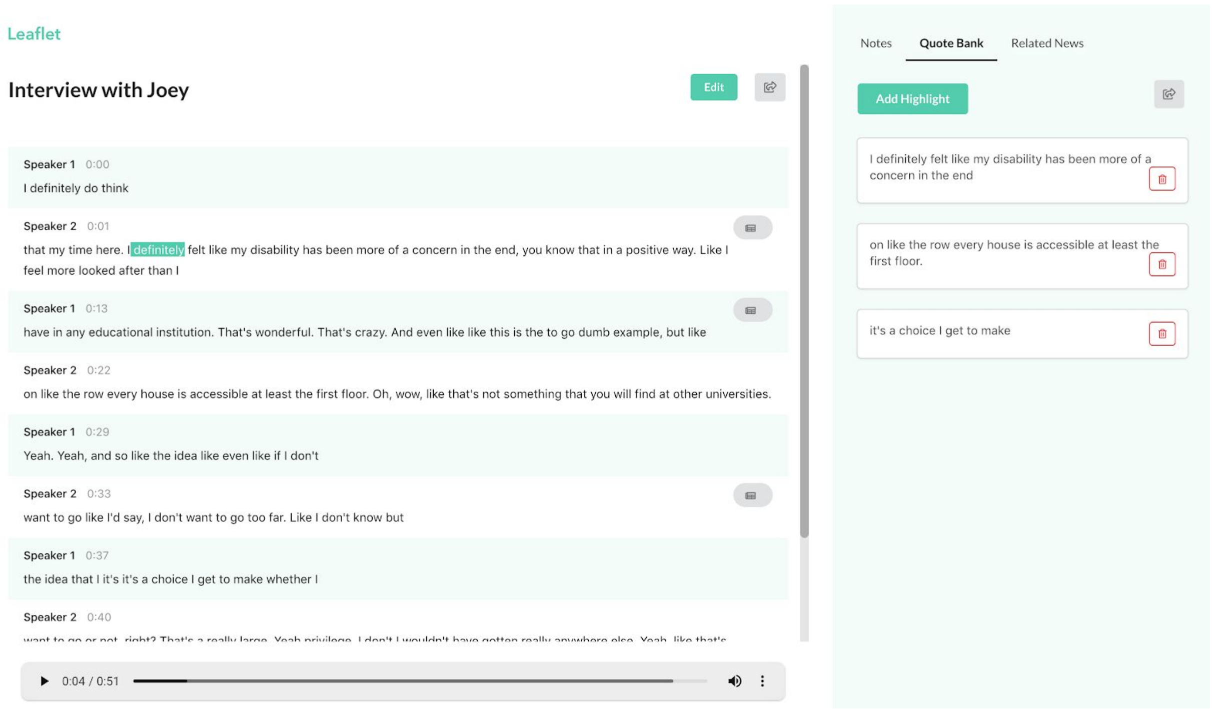
To implement automated transcription, we followed a similar-but-bumpier process. We read online comparisons of speech-to-text APIs, including a Medium article that crowned Google Speech-to-Text for video as the most accurate service. Speech-to-Text was challenging to get up and running due to its restrictions with audio file formatting and length. However, it produced stronger transcriptions than alternatives, including if we were to implement transcription from scratch. Natural language processing would be time intensive to do ourselves and companies like Google have the upside in running and testing their algorithms, which would be difficult to compete with in eight weeks. Additionally, Speech-to-Text provides useful features, such as speaker separation. We forged on, reused the Flask code from before, and voilà: Leaflet could take in your (under one minute) audio recording and print its transcript. With the data sent from the API and advice from Anh Truong, a computer science Ph.D. student, we were able to implement specialized audio features. These included transcript-audio matching (click on a word in the transcript and the corresponding audio plays), live highlighting (the word being played gets highlighted), visual separation of speakers, and more.

To implement text editing features, we used Draft.js, a rich-text editor framework for React. Draft.js allowed us to construct an editor with as many features as we needed. It gave us the flexibility to construct the design with Semantic UI, but also enabled seamless integration of features such as spell-check and keyboard shortcuts related to formatting.

In developing the related news feature, we researched APIs from a variety of news publications, such as the New York Times' API. Ultimately, we found the News API, which pulls articles from a variety of sources (over 30,000) to reflect the greater diversity of journalism. We implemented this as well the quote bank using React.

The Final Product

After weeks of furious developing, we introduce Leaflet: a modern transcription application designed to help journalists at every stage of their writing process.



Interview with Joey

Speaker 1 0:00
I definitely do think

Speaker 2 0:01
that my time here. I definitely felt like my disability has been more of a concern in the end, you know that in a positive way. Like I feel more looked after than I

Speaker 1 0:13
have in any educational institution. That's wonderful. That's crazy. And even like like this is the to go dumb example, but like

Speaker 2 0:22
on like the row every house is accessible at least the first floor. Oh, wow, like that's not something that you will find at other universities.

Speaker 1 0:29
Yeah. Yeah, and so like the idea like even like if I don't

Speaker 2 0:33
want to go like I'd say, I don't want to go too far. Like I don't know but

Speaker 1 0:37
the idea that I it's it's a choice I get to make whether I

Speaker 2 0:40
want to go or not, right? That's a really large. Yeah privilege. I don't I wouldn't have gotten really anywhere else. Yeah, like that's

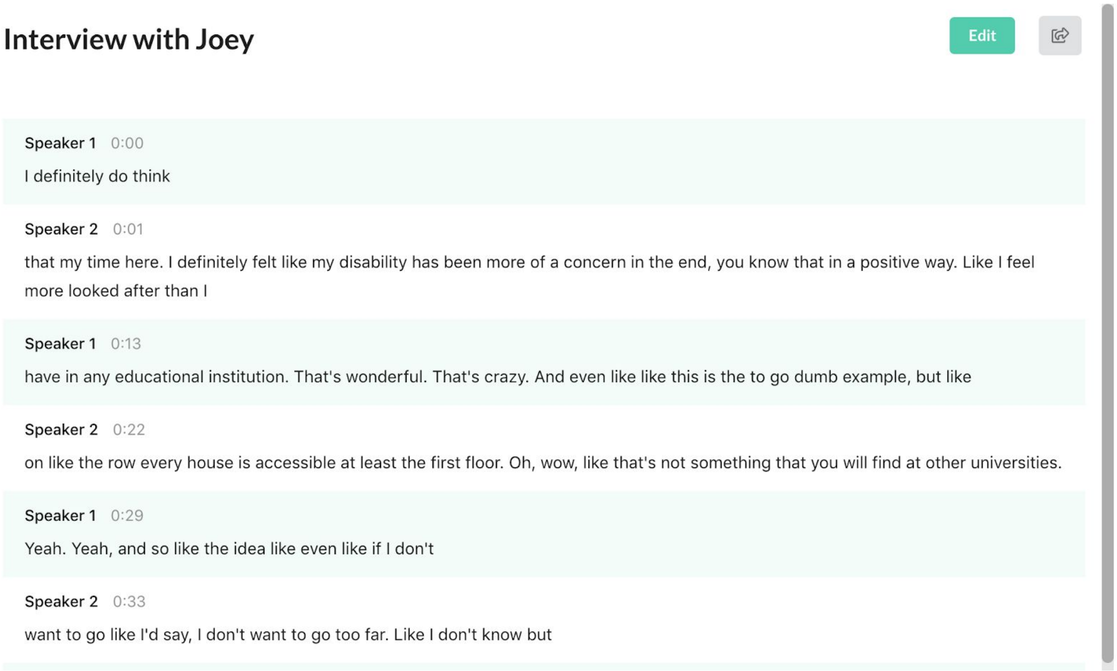
I definitely felt like my disability has been more of a concern in the end

on like the row every house is accessible at least the first floor.

it's a choice I get to make

Transcription

Journalists can upload an audio file and receive an automated transcript quickly.



Interview with Joey

Speaker 1 0:00
I definitely do think

Speaker 2 0:01
that my time here. I definitely felt like my disability has been more of a concern in the end, you know that in a positive way. Like I feel more looked after than I

Speaker 1 0:13
have in any educational institution. That's wonderful. That's crazy. And even like like this is the to go dumb example, but like

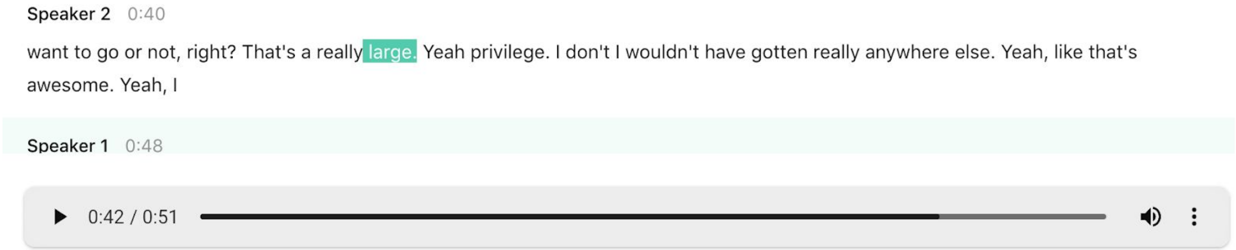
Speaker 2 0:22
on like the row every house is accessible at least the first floor. Oh, wow, like that's not something that you will find at other universities.

Speaker 1 0:29
Yeah. Yeah, and so like the idea like even like if I don't

Speaker 2 0:33
want to go like I'd say, I don't want to go too far. Like I don't know but

Playback Tools

Journalists can easily verify quotes by clicking on a word, which will play the audio from that word. An audio player and keyboard shortcuts also help journalists navigate through their recording.



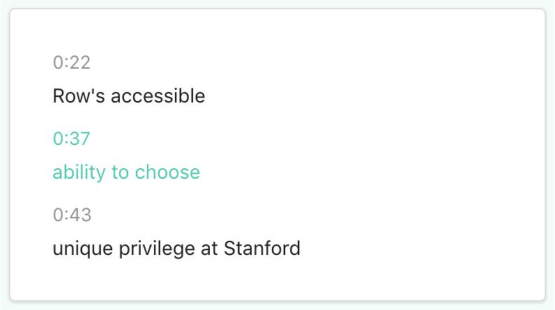
Speaker 2 0:40

want to go or not, right? That's a really large. Yeah privilege. I don't I wouldn't have gotten really anywhere else. Yeah, like that's awesome. Yeah, I

Speaker 1 0:48

Digital Notes

Journalists can upload an image of handwritten notes they took during the recording. Leaflet uses computer vision to digitally link each note to the corresponding section of the transcript.



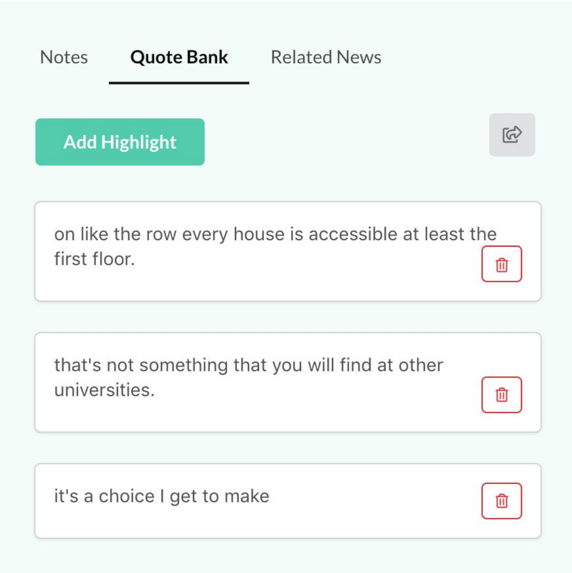
0:22
Row's accessible

0:37
ability to choose

0:43
unique privilege at Stanford

Quote Bank

Journalists can store important quotes within the transcript to use later when writing their story.



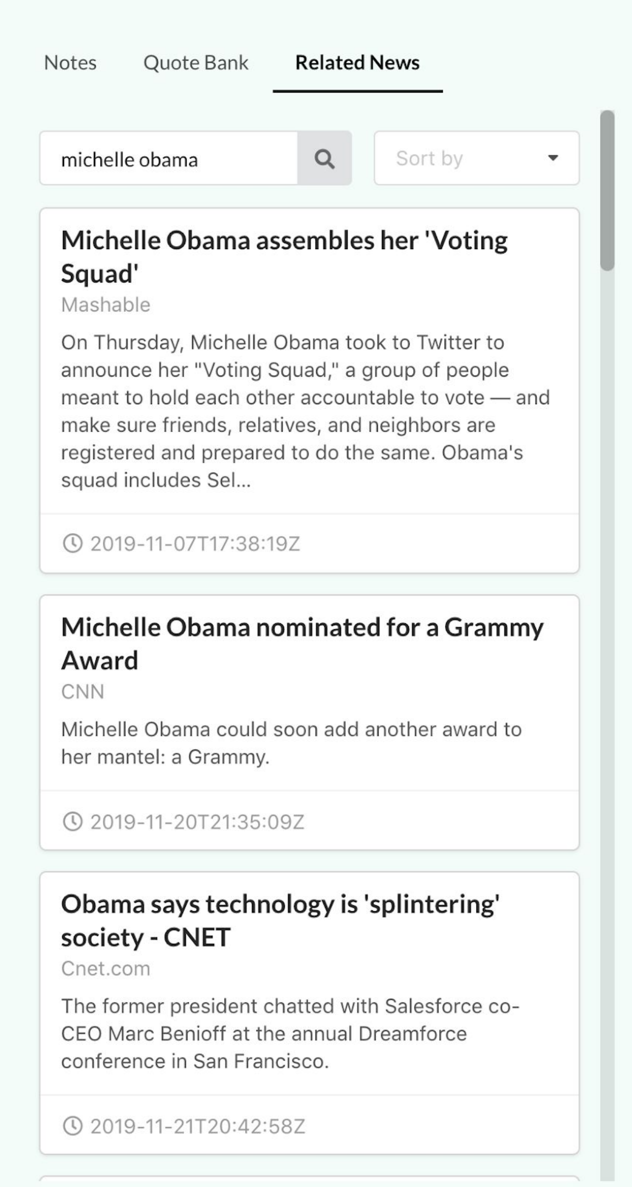
on like the row every house is accessible at least the first floor.

that's not something that you will find at other universities.

it's a choice I get to make

Related News

Journalists can quickly find out the latest news on any topic.



Michelle Obama assembles her 'Voting Squad'

Mashable

On Thursday, Michelle Obama took to Twitter to announce her "Voting Squad," a group of people meant to hold each other accountable to vote — and make sure friends, relatives, and neighbors are registered and prepared to do the same. Obama's squad includes Sel...

2019-11-07T17:38:19Z

Michelle Obama nominated for a Grammy Award

CNN

Michelle Obama could soon add another award to her mantel: a Grammy.

2019-11-20T21:35:09Z

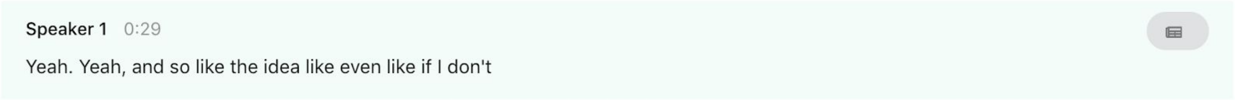
Obama says technology is 'splintering' society - CNET

Cnet.com

The former president chatted with Salesforce co-CEO Marc Benioff at the annual Dreamforce conference in San Francisco.

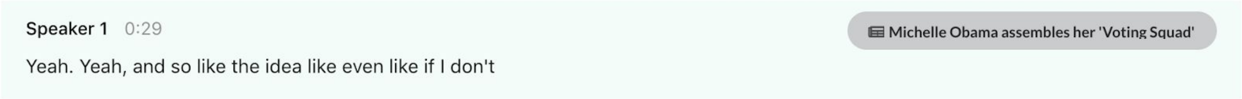
2019-11-21T20:42:58Z

They can drag and drop these news stories into the transcript to pin them for future reference.



Speaker 1 0:29

Yeah. Yeah, and so like the idea like even like if I don't

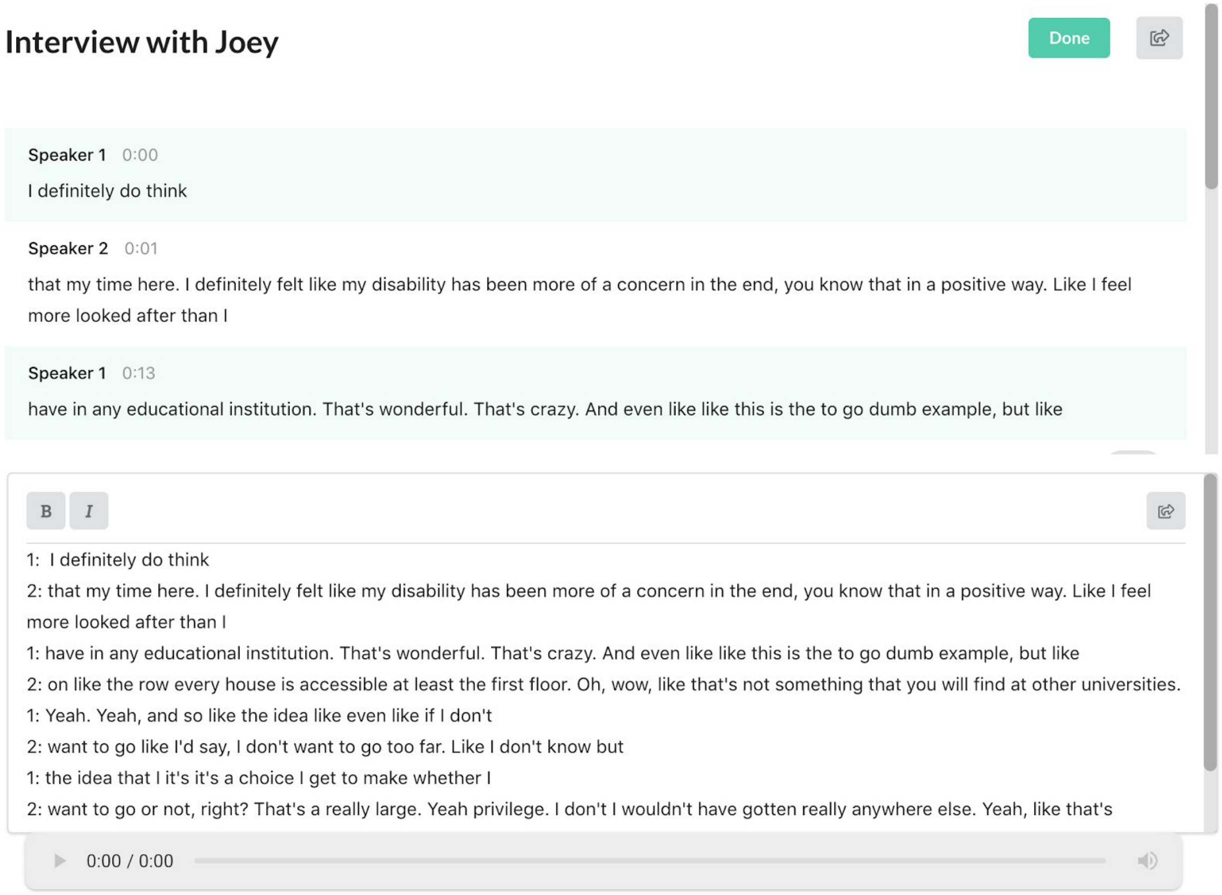


Speaker 1 0:29

Yeah. Yeah, and so like the idea like even like if I don't

Text Editor

Journalists can edit the transcript, rearrange quotes, and even begin writing their story.



Interview with Joey

Speaker 1 0:00
I definitely do think

Speaker 2 0:01
that my time here. I definitely felt like my disability has been more of a concern in the end, you know that in a positive way. Like I feel more looked after than I

Speaker 1 0:13
have in any educational institution. That's wonderful. That's crazy. And even like like this is the to go dumb example, but like

1: I definitely do think
2: that my time here. I definitely felt like my disability has been more of a concern in the end, you know that in a positive way. Like I feel more looked after than I
1: have in any educational institution. That's wonderful. That's crazy. And even like like this is the to go dumb example, but like
2: on like the row every house is accessible at least the first floor. Oh, wow, like that's not something that you will find at other universities.
1: Yeah. Yeah, and so like the idea like even like if I don't
2: want to go like I'd say, I don't want to go too far. Like I don't know but
1: the idea that I it's it's a choice I get to make whether I
2: want to go or not, right? That's a really large. Yeah privilege. I don't I wouldn't have gotten really anywhere else. Yeah, like that's

Future Direction

We created a number of features beyond sole transcription, which opens up new possibilities for journalists. We would love to gain a better understanding of use cases and how we can further integrate interactions across features. Beyond seeing this tool in action and iterating on feedback, there are also technical nuances we hope to improve. First, we would like the tool to accommodate for any number of speakers and length of recordings. We have a start, but there are currently limits on audio duration and accuracy of speaker detection that could hinder the journalist's experience. Second, we would like to push the capabilities for organizing notes. For example, currently the transcript is separated by speaker, but perhaps there could also be automatic sectioning based off of topic detection and keywords that the journalist marks as significant. These are considerations that perhaps at the beginning stages of Leaflet usage may not be evident, but a power-user with 20 pages of transcription may find these additional features valuable.

From a larger perspective, we see Leaflet as a means of bridging the world of audio and text. Audio is an exciting interface that we are increasingly familiarizing ourselves with, be it through podcasts or Amazon's Alexa. The increase in audio data presents an opportunity for more streamlined creation of auditory narratives. Today we have prototyping and collaboration tools for visual design, with products such as Figma, Sketch, and Webflow. Leaflet is our vision of audio experiences of the future. Audio should be a malleable medium that inspires creators rather than presenting a bottleneck.